Cisco Wholesale Voice Solution Overview

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:
Text Part Number: OL-1476-01

CONTENTS

Preface

The Cisco Wholesale Voice Solution is a service that provides trunk-level transport of global switched telephone traffic distributed by means of VoIP (voice over IP). The objective of this solution, or *set* of solutions, is to give service providers essential information about the required architecture design, network components, software features, functional areas, and provisioning methodologies needed to run a VoIP wholesale service.

This preface presents the following major topics:

- Document and Solution Release
- Audience
- Contents Overview
- Document Organization
- Related Documents
- Terms and Acronyms
- Obtaining Documentation
- Obtaining Technical Assistance

Document and Solution Release

This is the first release of this document, which covers Release 2.0(0) of the Cisco Wholesale Voice Solution. Release 1.0 contained a subset of the current features. Software upgrades or bug fixes to Release 2.0 will be indicated by 2.0(1), 2.0(2), and so on. As significant new features are added, the subsequent major releases will be indicated by 3.0(0), 4.0(0), and so on.

Audience

This document is part of a suite of documents for the following users:

- Candidate customers—who are able to take advantage of this solution and want to understand it
- Network operators/administrators—who have experience in telecommunications networks, protocols, and equipment, and a familiarity with data communications networks, protocols, and equipment
- Network designers—who have experience with telecommunications networks, protocols, and equipment, and experience with data communications networks, protocols, and equipment.

For the other documents related to the Cisco Wholesale Voice Solution, see Related Documents, page xi.

Contents Overview

The *Cisco Wholesale Voice Solution Overview* provides the following information:

- An introduction to the Cisco Wholesale Voice Solution, including the services it supports
- The solution's general architecture, network topologies, and components
- Deployment considerations, covering a variety of possible scenarios and issues that must be considered

This document does not provide detailed information on how to install and provision a given solution. Rather, it provides the background needed to understand the components, their interconnections, and key issues related to that solution. For the details of provisioning, refer to the *Cisco Wholesale Voice Solution Design and Implementation Guide*.

Note Figures and tables are listed in the Index, under "figures" and "tables," respectively.

Document Organization

The major sections of this document are as follows:

Section	Title	Description
Preface	Preface	Provides an overview of this document and lists related resources.
Chapter 1	Introduction	Introduces the Cisco Wholesale Voice Solution, listing candidate customers, benefits and features, and services.
Chapter 2	Solution Architecture	Discusses functional areas and call topologies.
Chapter 3	Solution Components	Lists components, both Cisco's and those of third parties, needed to implement Cisco Wholesale Voice Solutions. Lists both required and optional components. Includes hardware, IOS software, and third-party applications.
Chapter 4	Designing a Solution	Presents an approach to designing a solution for a given service.
Chapter 5	Service A: Minutes Aggregation and Resale (Including ASP Termination)	Presents key issues (dial plan, billing/settlement, security, prompting) as they relate to five basic templates for Service A.
Chapter 6	Service B: Card Services (Prepaid and Postpaid)	Presents key issues (dial plan, billing/settlement, security, prompting) as they relate to five basic templates for Service B.

Chapter 7	Deploying Service Options	Presents key issues (dial plan, billing/settlement, security, prompting) as they relate to several basic templates for Services A and B.
Appendix A	Open Settlements Protocol (OSP) Clearinghouse Solution	Provides a brief overview of OSP as it relates to clearinghouse applications.
Glossary	Glossary	Terms and acronyms used in this document

Related Documents

Refer to the following documents for detailed hardware and software installation and configuration information about the Cisco Wholesale Voice Solution:

- *Cisco Wholesale Voice Solution Design and Implementation Guide*

  The above provides links to online references for the many components of the solution.
- *Cisco Wholesale Voice Solution Release Notes*

Viewing Online Documents in Your Browser

As you click on links, the files you select may be added to the current document. When you close the file, you will be prompted to save the file. (You will not be able to save the file to a CD.) If you choose not to save the larger file that is created, click *No* when prompted to save the file. However, if you acquire documents that you want to save in a new file, you can save that file to another disk or drive with a new name of your own choosing. Set the following preferences within the Acrobat application to open weblinks in your browser, rather than within Acrobat.

You can obtain the latest version of Adobe Acrobat Reader at http://www.adobe.com.

Step 1 Select the browser you want to use.

a. From the Acrobat main menu, choose File > Preferences > Weblink. The Weblink Preferences window opens.

b. In the Weblink Preferences window, click Browse (or Select) and locate the browser you wish to use.

c. Then select Connection Type from the pull-down menu. Choose Standard if your browser is not listed.

d. Click OK to save your settings.

Step 1 Make sure that Acrobat opens weblinks in your browser.

a. From the Acrobat main menu, choose File > Preferences > Web Capture. The Web Capture Preferences window opens.

b. In the Web Capture Preferences Window, choose Open Weblinks: In Web Browser.

c. Click OK to save your settings.

Terms and Acronyms

For definitions of terms and acronyms used in the following chapters, refer to glossary at the end of this document.

For an online listing of internetworking terms and acronyms, refer to the following URL:

http://www.cisco.com/univercd/cc/td/doc/cisintwk/ita/index.htm

Obtaining Documentation

The following sections provide sources for obtaining documentation from Cisco Systems.

World Wide Web

You can access the most current Cisco documentation on the World Wide Web at the following sites:

- http://www.cisco.com
- http://www-china.cisco.com
- http://www-europe.cisco.com

Documentation CD-ROM

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

Ordering Documentation

Cisco documentation is available in the following ways:

- Registered Cisco Direct Customers can order Cisco Product documentation from the Networking Products MarketPlace:

  http://www.cisco.com/cgi-bin/order/order_root.pl

- Registered Cisco.com users can order the Documentation CD-ROM through the online Subscription Store:

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

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

Documentation Feedback

If you are reading Cisco product documentation on the World Wide Web, you can submit technical comments electronically. Click **Feedback** in the toolbar and select **Documentation**. After you complete the form, click **Submit** to send it to Cisco.

You can e-mail your comments to bug-doc@cisco.com.

To submit your comments by mail, for your convenience many documents contain a response card behind the front cover. Otherwise, you can mail your comments to the following address:

Cisco Systems, Inc.
Document Resource Connection
170 West Tasman Drive
San Jose, CA 95134-9883

We appreciate your comments.

Obtaining Technical Assistance

Cisco provides Cisco.com as a starting point for all technical assistance. Customers and partners can obtain documentation, troubleshooting tips, and sample configurations from online tools. For Cisco.com registered users, additional troubleshooting tools are available from the TAC website.

Cisco.com

Cisco.com (formerly known as Cisco Connection Online, or CCO) is the foundation of a suite of interactive, networked services that provide immediate, open access to Cisco information and resources at anytime, from anywhere in the world. This highly integrated Internet application is a powerful, easy-to-use tool for doing business with Cisco.

Cisco.com provides a broad range of features and services to help customers and partners streamline business processes and improve productivity. Through Cisco.com, you can find information about Cisco and our networking solutions, services, and programs. In addition, you can resolve technical issues with online technical support, download and test software packages, and order Cisco learning materials and merchandise. Valuable online skill assessment, training, and certification programs are also available.

Customers and partners can self-register on Cisco.com to obtain additional personalized information and services. Registered users can order products, check on the status of an order, access technical support, and view benefits specific to their relationships with Cisco.

To access Cisco.com, go to the following website:

http://www.cisco.com

Technical Assistance Center

The Cisco TAC website is available to all customers who need technical assistance with a Cisco product or technology that is under warranty or covered by a maintenance contract.

Contacting TAC by Using the Cisco TAC Website

If you have a priority level 3 (P3) or priority level 4 (P4) problem, contact TAC by going to the TAC website:

http://www.cisco.com/tac

P3 and P4 level problems are defined as follows:

- P3—Your network performance is degraded. Network functionality is noticeably impaired, but most business operations continue.
- P4—You need information or assistance on Cisco product capabilities, product installation, or basic product configuration.

In each of the above cases, use the Cisco TAC website to quickly find answers to your questions.

To register for Cisco.com, go to the following website:

http://www.cisco.com/register/

If you cannot resolve your technical issue by using the TAC online resources, Cisco.com registered users can open a case online by using the TAC Case Open tool at the following website:

http://www.cisco.com/tac/caseopen

Contacting TAC by Telephone

If you have a priority level 1(P1) or priority level 2 (P2) problem, contact TAC by telephone and immediately open a case. To obtain a directory of toll-free numbers for your country, go to the following website:

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

P1 and P2 level problems are defined as follows:

- P1—Your production network is down, causing a critical impact to business operations if service is not restored quickly. No workaround is available.
- P2—Your production network is severely degraded, affecting significant aspects of your business operations. No workaround is available.

Introduction

Wholesale Voice over IP

The investment by telephony and Internet carriers, both incumbents and startups, in voice over IP (VoIP) services is growing worldwide. A key reason for this growth is the simple economics of call transport. IP telephony has matured to the point where it can offer a significant and cost-effective alternative to traditional circuit-switch solutions. New advances in the areas of voice QoS, scaling, management, provisioning, and call routing have made it possible for service providers to offer managed IP backbone networks for the seamless transport of voice, fax, and data.

As wholesale carriers typically have low margins, wholesale VoIP providers can realize considerable profit by carrying large quantities of compressed traffic to selected areas—while optimizing their use of the network. Telephony on an IP backbone allows the service provider to carry more traffic over the same trunks by using advances in compression and call processing. It also offers conditional routing techniques, so that a service provider can adjust traffic patterns dynamically, in near real time, to ensure the best termination price. In addition, service vendors can realize many advantages in the rapidly emerging "enhanced services over IP" market segment. These services can adapt to the more open, scalable, and flexible call model of Open Packet Telephony (OPT).

Wholesale voice is a growth market, with service providers building new capacity and launching new services. The primary wholesale service is long-distance transport and aggregation, with the key advantage that country-specific features and domestic calling regulations are not required. Principal beneficiaries are developing countries, where in many cases the quality of VoIP is superior to that of traditional PSTN service.

This chapter presents the following major topics:

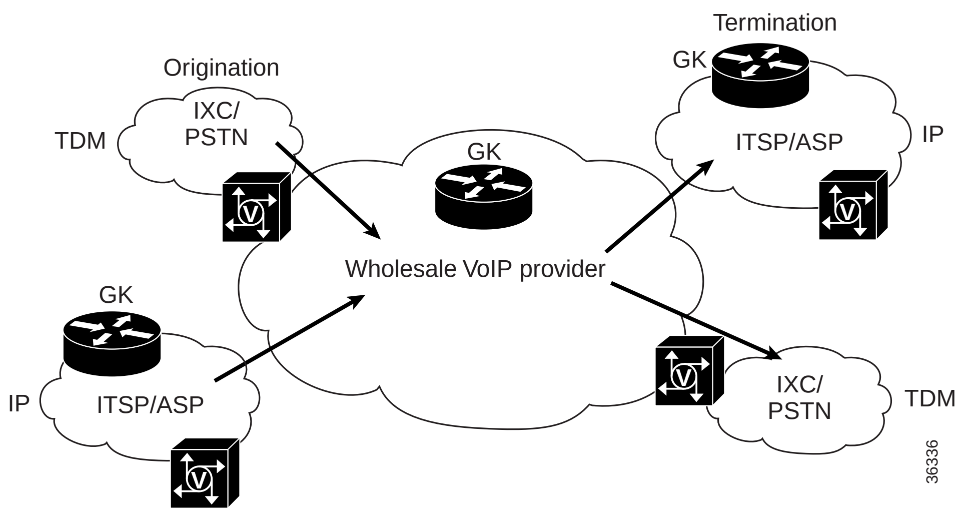
- The Cisco Wholesale Voice Solution
- Services Supported
- Architecture

The Cisco Wholesale Voice Solution

The Cisco Wholesale Voice Solution provides service providers (SPs) with the required architecture design, network components, software features, functional areas, and provisioning methodologies needed to run a VoIP wholesale service. With an understanding of the concepts underlying the architecture, including interconnect topologies, components, and a variety of important issues that must be considered, the SP can then deploy options from a set of configuration templates. The result is a wholesale network that allows the SP to sell unbranded voice services to retailers, such as Internet telephony service providers (ITSPs), application service providers (ASPs), interexchange carriers (IXCs), or Post Telephone and Telegraph administrations (PTTs).

Central to the delivery of wholesale voice services are voice points of presence (POPs), which are interconnected to other service providers. The specific recommended components and design methods are determined by the type of interconnection or "call topology" that the wholesale SP is supporting. These call topologies are used to build a set of deployment templates for an SP to enable wholesale applications. Figure 1-1 illustrates the interconnection possibilities that a wholesale VoIP provider must accommodate.

Figure 1-1 Possible Interconnect Scenarios



The Cisco Wholesale Voice Solution is a *set* of solutions and network design and configuration templates that provide trunk-level transport of global switched telephone traffic distributed over VoIP. Calls originate in the PSTN, are routed through IXCs, and are handed off to a wholesale VoIP carrier for transport. To the end user, the service looks like any other long-distance call, except that it is less expensive. To the originating long-distance carrier, the wholesale carrier is only one of a number of termination options.

By using OPT distributed architectures, the Cisco Wholesale Voice Solution maintains separate call control, connection control, and transport planes. At the heart of the solution are Cisco gateways (GWs), gatekeepers (GKs), and directory gatekeepers (DGKs), as well as an IP backbone. This solution will provide other network providers with connectivity between basic telephone areas and international routes. (Local residential services and features will not be provided.) The remaining components of the solution are third-party shared services that will vary with each application—such as settlement servers, billing servers, and AAA (authentication, authorization, and accounting) servers, among others.

Benefits and Features

The Cisco Wholesale Voice Solution provides the following benefits:

- Voice quality that is comparable to that of the PSTN
- A cost-effective, reliable VoIP network infrastructure
- Support for least-cost routing and other enhanced call-routing methods

- Intercarrier call authorization and accounting (peer to peer)
- Support for intercarrier clearinghouse and settlement services
- Support for local, national, and international dial plans
- Connectivity with the PSTN over carrier interfaces
- Connectivity with other VoIP service providers and other vendors' VoIP equipment
- A world-wide network of other VoIP service providers interested in interconnecting

Table 1-1 briefly lists features that are provided with the Cisco Wholesale Voice Solution.

Table 1-1 Features of the Cisco Wholesale Voice Solution

Feature	Purpose
Broad interoperability and standards support	Support wide range of H.323 intercarrier call-routing requirements and options
Call accounting and authorization applications	Benefit from Cisco Ecosystem Partners such as MIND CTI and Belle Systems
Flexible call-routing options	Support open API for call routing on top of Cisco gatekeeper
Flexible PSTN signaling options	Support widely accepted E1/R2 signaling for European, South American, and Asian markets; T1 for North American markets; and SS7 signaling for medium and larger POPs
High-performance embedded-system H.323 gatekeepers	Provide VoIP interconnect, interoperability with other vendors
Industry-leading voice quality	Eliminate distinction between low-cost VoIP service and high-cost PSTN service
Inexpensive, integrated programmable IVR (interactive voice response)	Support prepaid calling services, enabling wholesale carrier to add higher-margin subscriber services to the network
Multiple H.323 gateway options	Provide appropriate platform support for both small and medium-size POPS
Partnerships with major settlement providers	Benefit from Cisco Ecosystem Partners such as TransNexus

Candidate Customers

The following customers will be the prime beneficiaries of the Cisco Wholesale Voice Solution:

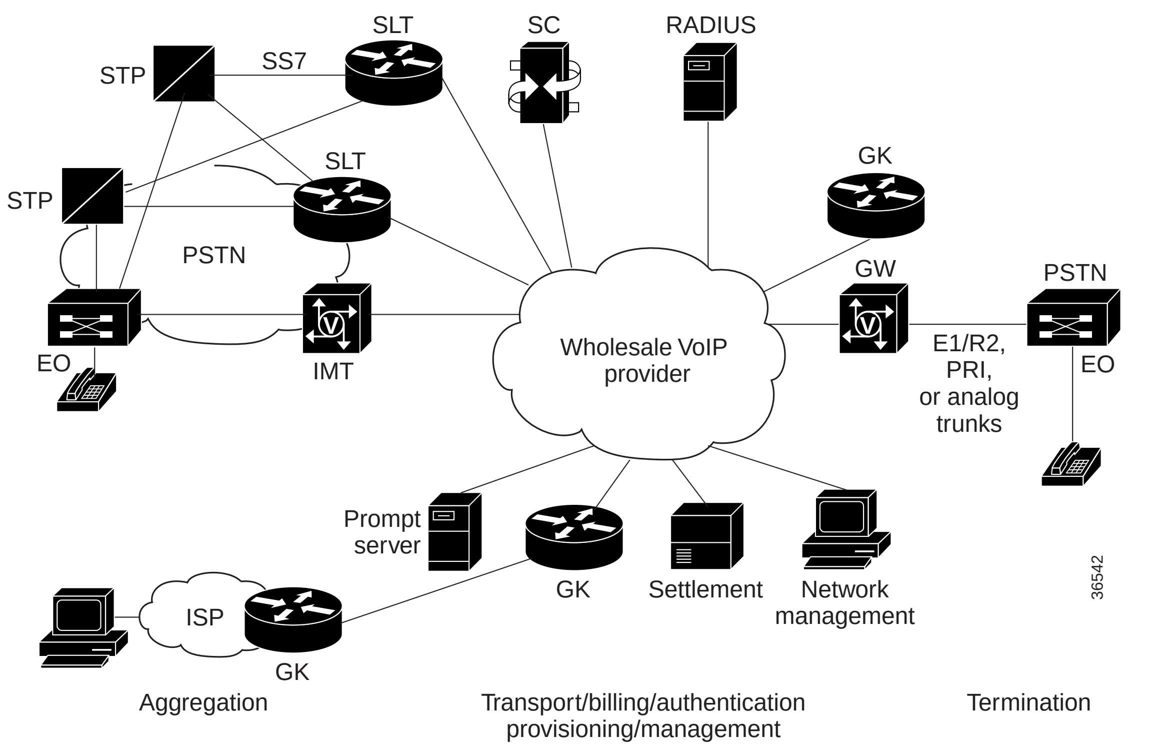
- Wholesale voice resellers
- Internet telephony service providers (ITSPs) and wholesale VoIP operators
- Interexchange carriers (IXCs)
- International carriers
- Local exchange carriers (Less) with wholesale agreements
- So-called "other licensed operators," or OLOs (such as mobile operators)
- Network service providers terminating calls for ISP-originated services

- Traditional voice carriers using VoIP for interconnecting with other carriers to bypass congested routes, targeting specific international routes
- Traditional carriers migrating existing wholesale business from TDM to VoIP
- VoIP for interswitch tandem bypass between consortia members' international switches across international boundaries (or within a single carrier)
- Switch-based resellers integrating VoIP routes, including callback and reoriginators implementing VoIP for international toll bypass and debit
- New carriers building business cases based on VoIP cost savings
- Wireless/competitive carriers building VoIP "backbone" networks, extending this for international interconnect
- New or existing service providers

Services Supported

Figure 1-2 summarizes all of the components that may be needed to provide Cisco Wholesale Voice Solution services. The various scenarios discussed later in this document focus on the specific components needed for each scenario.

Figure 1-2 High-Level View of End-to-End Service Possibilities



Naturally, not all services will use all components. The service models presented below can overlap with respect to a given customer. For example, one customer could be both a minutes aggregator/terminator as well as a clearinghouse. Conversely, although a clearinghouse is most often viewed as part of a solution, a single customer providing clearinghouse services is also a candidate for the Cisco Wholesale Voice Solution. A key feature of the Cisco Wholesale Voice Solution is its ability to support various mixes of the following services to suit the needs of a single or multiple (partnering) service providers.

The Cisco Wholesale Voice Solution supports the following services:

- Service A: Minutes Aggregation and Resale (Including ASP Termination)
- Service B: Calling Card Services (Prepaid and Postpaid)

These are introduced below.

Service A: Minutes Aggregation and Resale (Including ASP Termination)

This service allows wholesale network providers to collect traffic from multiple originating providers, then aggregate and deliver it to termination providers they select. These providers may include target greenfields, resellers, dial-around callback operators, and international ISPs.

The Cisco Wholesale Voice Solution supports the originating carrier, who profits up front by handing calls over to a VoIP wholesaler. Termination settlement rates are generally lower than PSTN termination rates—the key reason why long-distance carriers will choose a VoIP carrier for termination. Furthermore, termination bandwidth is often available over VoIP to countries where PSTN termination is unavailable (for example, because of congested international gateway facilities).

Voice, modem, and fax calls are supported. The interfaces supported are SS7, T1/E1 PRI, E&M, and R2. Average call success rate is as good as or better than that provided by PSTN carriers. Furthermore, voice quality, including echo cancellation, is uncompromised.

Key features of this service include the following:

- H.323 VoIP interconnect, using standards-based H.323 implementation
- Gatekeeper LRQ forwarding for call routing

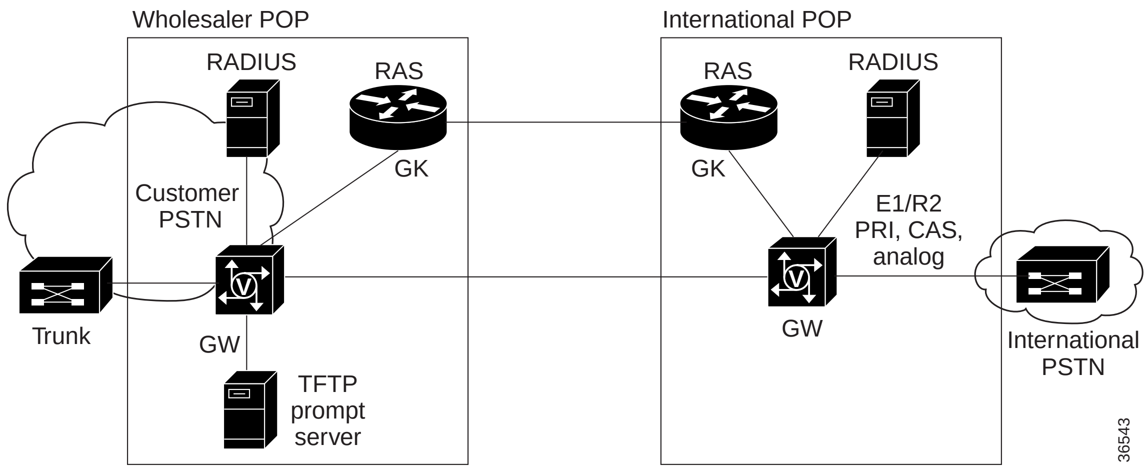
As part of Service A, the Cisco Wholesale Voice Solution supports termination services for ASP-originated calls. In this carrier-to-carrier service, the ASP originates the call, often over an Internet-enabled PC-telephony application, or through a PSTN portal for cellular phone callers. The ASP provides precall services, such as content delivery (prerecorded messages, voice mail, private number dialing) or supervision-related services (“find me/follow me”). The ASP then needs to hand any long-distance calls off to a wholesale carrier for termination by the PSTN.

For the key issues related to this service for various call topologies, refer to Chapter 5, “Service A: Minutes Aggregation and Resale (Including ASP Termination).”

Service B: Calling Card Services (Prepaid and Postpaid)

The Cisco Wholesale Voice Solution supports both prepaid and postpaid calling-card services, as described below. These are both subscriber services. Figure 1-3 illustrates, as an example, the components and signaling required to provide calling card services for international calls.

Figure 1-3 Providing Calling-Card Services for International Calls



Prepaid

In the prepaid service, a wholesale VoIP carrier can host prepaid services for multiple service providers on their infrastructure. In addition, most prepaid service providers use VoIP wholesalers to terminate long-distance calls that are placed by prepaid subscribers.

Using the interactive voice response (IVR) feature in the Cisco VoIP gateways, and real-time authorization and call accounting systems provided byCisco Ecosystem Partners, service providers can offer this service over a VoIP network and lower the cost and deployment time of calling-card services.

Note Cisco has identified leading solution providers with which to partner in offering end-to-end benefits. These providers are collectively known as the Cisco New World Ecosystem. For more information about the Ecosystem Partner community, visit http://www.cisco.com/public/Partner_root.shtml.

Postpaid

Like the prepaid service, postpaid service can be hosted by a wholesale VoIP carrier. An example is basic calling that is accessed by the 800 prefix, a calling card number, or a PIN. Postpaid is similar to the prepaid service, except that with postpaid the authorization is not tied to call rating. Consequently, call rating does not have to happen in real time, and there may be more partner billing-system options that perform adequately at scale. After calls are made, a billing system contracted by the company charges the carrier.

For the key issues related to this service for various call topologies, refer to Chapter 6, "Service B: Card Services (Prepaid and Postpaid)."

Clearinghouse Services

Where multiple partners are involved, as will often be the case, the above services may require the assistance of clearinghouse services for billing, settlement, and invoicing. These services can be based on Open Settlements Protocol (OSP), or on other methods of mediation and settlement, including AAA/RADIUS. Where the services are OSP based, the Cisco Wholesale Voice Solution supports call termination agreements through support for OSP in Cisco devices.

Note For more information about OSP, see Appendix A, "Open Settlements Protocol (OSP) Clearinghouse Solution."

Service Options

The following enhancements are provided to support the services offered by Cisco Wholesale Voice Solution:

- Limited Egress Carrier-Sensitive Routing
- Interconnect to Clarent-Based Clearinghouses

Limited Egress Carrier-Sensitive Routing

As an enhancement to simple carrier-interconnect applications, the Cisco Wholesale Voice Solution makes it possible to route a call to different destination carriers. The wholesaler has the same considerations as with simple carrier-interconnect models, but with slightly increased call-routing responsibilities. For more information, refer to Limited Egress Carrier-Sensitive Routing in Chapter 7, "Deploying Service Options."

Interconnect to Clarent-Based Clearinghouses

Cisco-based wholesaler VoIP networks can interconnect with a Clarent-based service provider (http://www.clarent.com). For more information, refer to Interconnect to Clarent-Based Clearinghouses in Chapter 7, "Deploying Service Options."

Architecture

The Cisco Wholesale Voice Solution architecture builds upon existing Cisco products to provide services that will run over the service provider's VoIP infrastructure. To understand that architecture, including functional areas and components, proceed to Chapter 2, "Solution Architecture."

Solution Architecture

This chapter introduces the functional areas, interconnection types, and components that make up the architecture of the Cisco Wholesale Voice Solution. This provides a foundation for understanding the deployment of various solutions, and introduces a variety of issues that must be taken into consideration for each solution.

This chapter presents the following major topics:

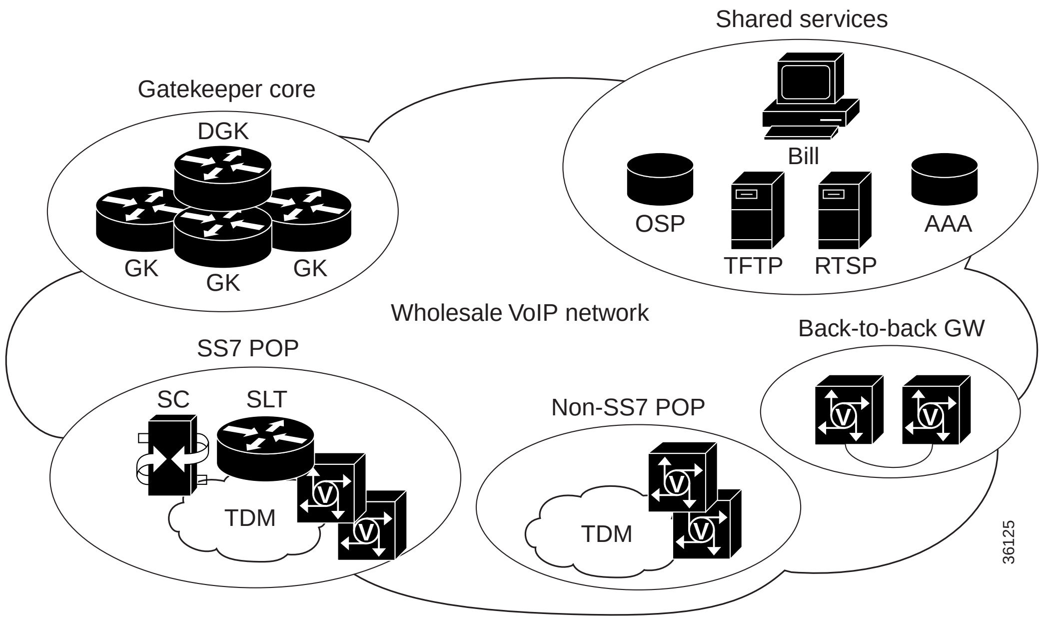
- Functional Areas
- Call Topologies and SP Interconnection Methods
- Deployment Templates

Functional Areas

The wholesale VoIP provider may need to accommodate a variety of factors in various combinations, depending on the SP's partners. These factors include, but are not limited to, interconnection methods, billing services and settlement, call control, IVR options, and network management. To support this variability, the Cisco Wholesale Voice Solution encompasses five functional areas designed to respond to those requirements.

Figure 2-1 depicts these functional areas with respect to the wholesale VoIP network. Each of the areas, including interconnection methods and related issues, will be discussed later.

Figure 2-1 Functional Areas of the Cisco Wholesale Voice Solution

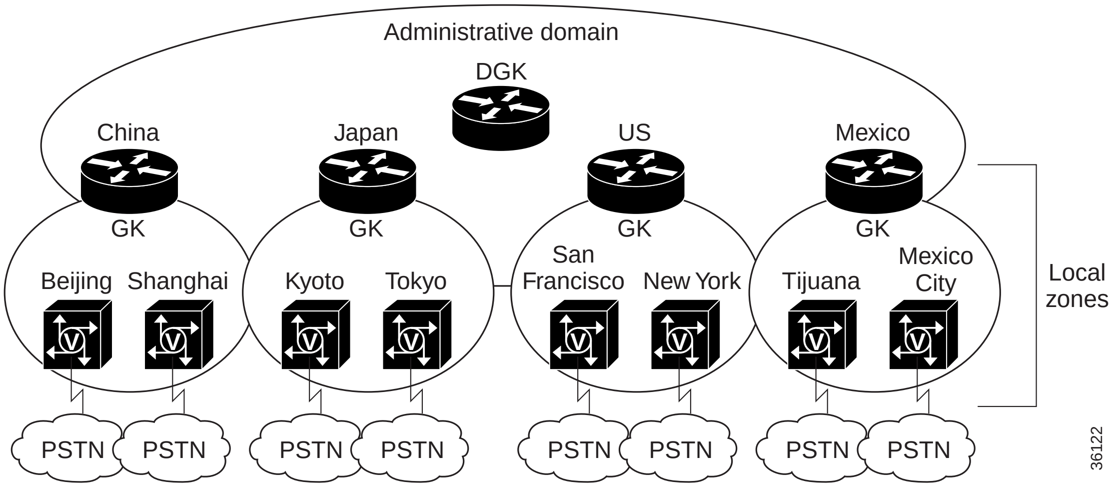


Each functional area has responsibilities and deployment considerations that are influenced by the application and call topology. For example, in the case of simple carrier interconnect with an ITSP, the wholesale provider must consider the configurations and components of each functional area throughout the network, such as POPs, GKs, and the shared support services needed to support routing, billing, settlement, and security options. The functional areas are discussed below.

Gatekeeper Core

The GK functional area, illustrated in Figure 2-2, consists of Cisco GK, Cisco DGK, and optional Ecosystem Partner GK platforms; it is the foundation of a large-scale H.323 network design. GKs enable a network to scale in growth, performance, and dial plan administration. They also enable some intelligent call routing and fault tolerance features. In addition, GKs support interactions with shared services and provide GK-based interconnections with other providers if the application demands.

Figure 2-2 Role of DGK and GKs in the Gatekeeper Core



Note Note: Cisco 3660s and 7200s are used as GK or DGK platforms.

Issues related to the gatekeeper core are discussed below.

Scaling Network Size

Large H.323 VoIP networks are segmented into different regional zones, each managed by a GK. Segmentation is based upon several factors, such as the desired call throughput (for example, BHCA), the dial plan, and the number of active endpoints.

As network coverage and capacity grows, the service provider may expand by simply adding new GWs or POPs to GKs until performance limitations for the GK platform are met. At that point, the service provider may expand by simply adding new GKs. Traffic is routed between GK zones by means of LRQ/LCF RAS messages.

Scaling Dial Plans

As more GKs are added to the network, inter-GK routing configurations increase dramatically. The smallest change to the dial plan requires configuration changes to all GKs in the network. At medium scales (that is, where the number of zones is relatively small), these changes can be managed by having a single dial plan that is downloaded through TFTP to all of the GKs within the wholesaler's administrative domain. As scale increases, the number of zones and the rate of dial plan updating increases. In this case, rather than burdening every GK with routing information for the entire network, a DGK is used to isolate and alleviate dial plan provisioning. For more information about dial plan provisioning across functional areas, see Basic Dial Plan, page 4-7.

Fault Tolerance

Cisco GKs and DGKs can be designed to enable redundancy in the dial plan. At the edge, GWs at each POP are configured to support registration with an alternate GK in the event the primary GK fails. In the core, GKs are configured to support sequential LRQ messages to provide redundant paths to alternate DGKs (ALTDGKs) and also accommodate local zone-fragmentation conditions. At the DGK level, both sequential LRQs to accommodate zone fragmentation and HSRP (Hot Standby Routing Protocol) are configured to provide redundancy at the highest layer. For more information about fault-tolerant architectures across functional areas, see Fault Tolerance, page 4-8.

DGK-Based IP Interconnect

Wholesale providers have the option to interconnect routes with other service providers by using a DGK. In this model, the wholesaler and the interconnect partner configure the DGKs to exchange LRQ RAS messages between their administrative domains to resolve call routing for desired prefixes. Sequential LRQs may be implemented on the DGK to support limited-egress CSR applications. Back-to-back GWs may be used to support IP-to-IP call topologies. Specific details are discussed for the deployment templates presented later in this document.

Security

To validate whether or not a call was originated from a valid endpoint, Cisco GWs and GKs may optionally implement access lists to provide secure gateway registration. In order to support this, GKs must be configured to interact with an AAA server. For more information about security across different functional areas, see Security, page 4-9.

Network Management

GKs must be enabled to support SNMP community strings, so that external management platforms such as CiscoWorks2000 Voice Manager (CVM) version 2.0.2 and Cisco Info Center (CIC) can provision networks, access reporting information, and receive traps through SNMP. Cisco Voice Manager is part of the CiscoWorks2000 suite.

TFTP Server Access

If the wholesaler desires, the GK may be configured to support the remote downloading of, for example, software images, configurations, TCL scripts, and audio files through a TFTP server.

Shared Support Services

Shared support services are central resources that enable network applications in the areas of card services, call routing, billing, settlement, security, and network management. The primary elements that enable these applications are AAA servers, billing systems, OSP servers, EMS platforms, NMS platforms, and TFTP servers.

Issues related to shared support services are discussed below.

Billing

An AAA server collects usage records directly from the GWs. Alternatively, an OSP server may also collect usage records for only interdomain calls. The details of billing implementations vary according to the application enabled.

Security

Complex access lists may be provisioned on the GWs to implement security functions. In this case, a TFTP server may be used as a central location for storing, administering, and uploading GW configurations. (However, this method is highly unscalable.)

An OSP server supplies security functions for OSP interconnect methods. For more information about implementing security across functional areas, see Security, page 4-9.

Note The Cisco Wholesale Voice Solution supports Cisco H.235 security for *registration only*. For reasons of AAA latency, H.235 security is not supported on a per-call basis. When per-call security is required, Cisco recommends that you use access lists.

Call Routing

For OSP-based interconnect scenarios, an OSP server handles call routing functions, as well as some complementary provisioning, on the OSP GW dial peers. Additionally, it is possible for an external server to provide enhanced call routing functions by interfacing with Cisco GKs and DGKs through GKTMP (GateKeeper Transaction Management Protocol). For a discussion of the impact of call routing on the dial plan, see Basic Dial Plan, page 4-7.

Note GKTMP applications are not specified for the Cisco Wholesale Voice Solution, nor are they tested.

Network Management

Standard SNMP NMS platforms (for example, HP OpenView™) may be used to provide generic SNMP management functions. CVM 2.0.2 provides SNMP-based statistics collection, as well as a very limited dial plan and component provisioning tool. Reports may be generated by means of Ecosystem Partner reporting engines that integrate with CVM 2.0.2. CIC (Cisco Info Center) may be used if fault management is desired. Additionally, Cisco's Internetwork Performance Monitor (IPM) may be used to monitor network QoS.

Note CVM 2.0.2 is recommended as the NMS platform. The Cisco Wholesale Voice Solution recognizes Trinagy (http://www.trinagy.com) as a vendor of report-generating applications.

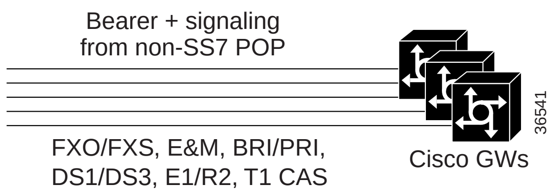
Remote Download

A remote TFTP server may be used to store IVR prompts, TCL scripts, software images, and configurations for download.

Non–SS7-Based POPs

Wholesale service provider networks consist of POPs that house gateways to transport voice traffic between TDM and IP networks. POPs are active components in call topologies 1, 2, and 3, as discussed in IP Interconnect Variations, page 4-3. Non-SS7 based POPs, which tend to be smaller than their SS7-based counterparts, receive signaling from the TDM network on the same physical interface that supports bearer traffic. There may be a logical separation between signaling and bearer traffic within the interface, as there is with ISDN. Actual gateway platforms used at these POPs will depend upon the signaling type offered by the TDM interconnect. Figure 2-3 illustrates non–SS7-based POP signaling. Interfaces supported through in-band signaling are FXO/FXS, E&M, BRI/PRI, DS1/DS3, E1/R2, and T1 CAS.

Figure 2-3 Non–SS7-Based POP Signaling



Note The Cisco AS5800 is not currently supported in the Cisco Wholesale Voice Solution.

In addition to the physical interface and signaling variations, there are also a number of platform-independent software features and functions that must be enabled on the POP GWs to support an application. These include POP size, dial plan, fault-tolerance, security, billing, network management, and bearer transport responsibilities.

Issues related to non-SS7-based POPs are discussed below.

Signaling Types

The signaling types can vary greatly and can include analog FXO, analog E&M, , DS1 interfaces with E1/R2 variations or T1 CAS variations, and may include DS3 interfaces on the upper boundary.

Low-density analog interfaces generally discourage carrier interconnects, so calls that ingress the POP will almost always be for card services. Calls that egress the POP are re-originated into the PSTN, usually to bypass PTT interconnect tariffs. Generally, DS1 and DS3 interfaces either provide card services or interconnect wholesale systems to their customers.

Note Cisco 3600 series, Cisco AS5300, Cisco AS5350, and Cisco AS5400 gateways with the appropriate analog or digital interfaces are used. For the latest information refer to the most current *Cisco Wholesale Voice Solution Release Notes*.

Size

Additional considerations arise with small-scale POPs. The hardware footprint of the equipment, as well as the amount of nonbearer traffic, must be minimized, because the IP network bandwidth coming into the POP from the wholesaler is likely to be extremely low (assume sub-E1 bandwidth).

Dial Plan

Dial plans encompass more than one functional area, and their responsibilities are distributed among GWs, GKs, and DGKs. The GWs have to deal with the local POP portion of the dial plan. This includes provisioning the required dial peers, translation rules, and RAI thresholds. For more information refer to Basic Dial Plan, page 4-7.

Billing

For reasons of performance and accuracy, Cisco recommends that billing be done from the GW whenever possible. The wholesaler must configure the Cisco GWs to interact with shared AAA services to support billing and optional debit card applications.

Fault Tolerance

If the wholesaler desires, the GW may be configured to support an alternate GK with which it will register should the primary GK fail. To support this concept fully, this implies a related configuration in the GK functional area. For more information about fault tolerance and its requirements across functional areas, see Fault Tolerance, page 4-8.

Security

Different security options are available to a wholesaler. To support security, GWs may be configured with complex access lists. For OSP-based interconnect scenarios, the GWs must be provisioned to interact with the OSP server to support OSP security options. For more information about implementing security across functional areas, see Security, page 4-9.

Network Management

GWs must be enabled to support SNMP community strings, so that external management platforms such as CVM 2.0.2 and CIC can provision, access reporting information, and receive traps through SNMP.

Transparent Bearer Transport

Unless the wholesaler has previously agreed to limit the types of calls exchanged between other carriers, the wholesaler providing interconnect may receive traffic of any bearer type. The wholesaler GWs must be able to pass voice, real-time fax, and modem traffic transparently across the VoIP network.

TFTP Server Access

If the wholesaler desires, the GW may be configured to support the remote downloading of prompts, software images, and configurations through a TFTP server.

SS7-Based POPs

These POPs have generally the same deployment considerations as the non–SS7-based POPs with respect to billing, security, network management, transparent bearer transport, and TFTP servers. However, here there are additional considerations related to SS7 interconnect signaling. Added considerations also appear with respect to POP size, dial plan responsibilities, and fault tolerance.

Issues related to SS7-based POPs are discussed below.

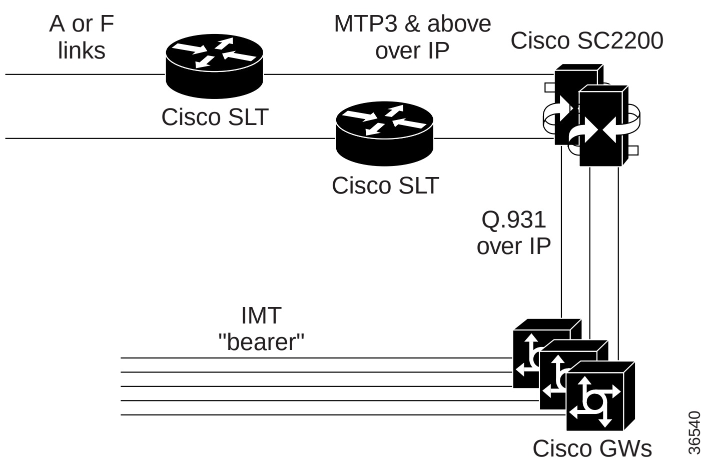
Signaling

SS7-based POPs tend to be larger than their non-SS7 counterparts, and consist of DS1 and DS3 IMTs to the GWs. PSTN-side call control is provided by using Q.931 backhaul from Cisco SC2200s running Cisco SS7 Interconnect for Voice Gateways to Cisco AS5300 series GWs. POPs may optionally support Cisco SLTs (Signaling Link Terminals) to terminate SS7 signaling on behalf of the Cisco SC2200 Signaling Controller.

Note Cisco SC2200s running Cisco SS7 Interconnect for Voice Gateways are recommended, as are Cisco 2611 SLTs. A Cisco SC2200 *node* is defined as SC2200 hosts and SLTs in redundant (fault tolerant) pairs. Cisco AS5300, Cisco AS5350, and Cisco AS5400 routers may be used as GWs in these configurations. Cisco 3600 series routers are not used where SS7 signaling is required.

Figure 2-4 illustrates the signaling used in an SS7 POP, showing the relationship among Cisco SC2200 hosts, Cisco SC2611 SLTs, and Cisco GWs.

Figure 2-4 SS7-Based POP Signaling



Dial Plan

For SS7-based POPs, the wholesaler has the option of performing number modification in either the GW, the Cisco SC2200 signaling controller, or both. The Cisco SC2200 allows digits in the called or calling party number fields to be added, deleted, or modified. It is also possible to modify the nature of address (NOA), perform black listing and white listing, and do AIN triggering. In addition to normal H.323

configurations, the GW must be provisioned with an RLM (Redundant Link Management) group in order to interface with the Cisco SC2200. Once the Cisco SC2200 and GW are provisioned to interface with each other, the rest of the H.323 dial plan remains the same as discussed in Basic Dial Plan, page 4-7.

Fault Tolerance

The GW may support a backup Cisco SC2200 in case the primary SC2200 fails. It may take up to 3 seconds for the GW to detect and fail over to the other Cisco SC2200. During this time, any new calls will not be processed. Furthermore, any calls that were in the process of being set up are also lost. Calls that are active at the point of failover, however, remain in progress.

GWs may support the same H.323 fault-tolerance conditions for GK failover as discussed in Fault Tolerance, page 4-8.

Note Cisco recommends that you provide redundant Cisco SC2200 hosts and Cisco SLTs in SS7 applications.

Back-to-Back GW

The back-to-back GW is a special component used for a variety of applications—particularly in billing (see Billing, page 2-4), by establishing a point in the call-signaling path from which to bill. GWs are deployed as a pair in a back-to-back TDM T1 CAS, ISDN PRI, or E1/R2 trunk, single-stage dialing configuration. Depending on the application, back-to-back GWs may function as unidirectional or bidirectional call devices.

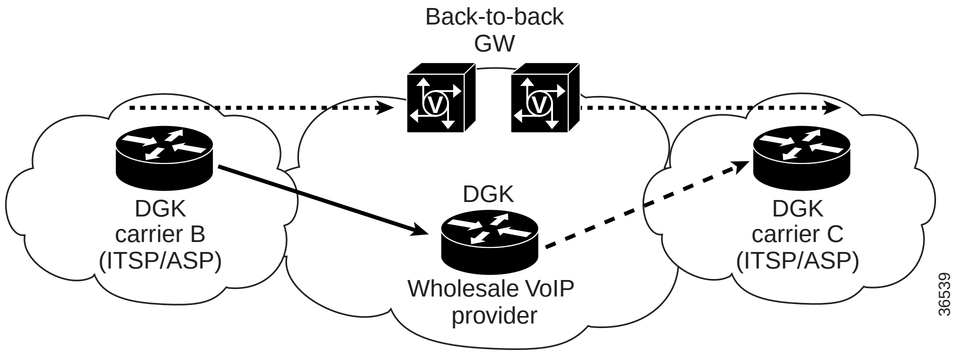
Note Cisco recommends that you use ISDN T1/E1 signaling between the GWs in a back-to-back pair.

For example, in an IVR application, the back-to-back GW has a dedicated half that receives all calls while the other half is dedicated to originating calls. In contrast, for an OSP interconnect zone application, the back-to-back GW may process calls in both directions, although each GW is responsible for separate protocols.

Although consisting of two GWs, the term “back-to-back GW” is most frequently used to refer to both units as an integrated pair. For unidirectional applications, to provide added clarity in discussing back-to-back GW pairs, we refer to the individual GWs in a pair as an *ingress* (inbound) VoIP GW and an *egress* (outbound) VoIP GW with respect to the call direction. For bidirectional applications, we generally refer to the GW by the protocol it supports.

Figure 2-5 illustrates the relationship of the back-to-back GW to the wholesaler and an ingress and egress carrier.

Figure 2-5 Relationship of the Back-to-Back GW to Wholesaler and Carriers



The GW pair allows wholesalers to insert themselves into the call-signaling path in IP-to-IP interconnect call topologies. This allows them to generate usage records by means of AAA, interconnect with Clarent-based and OSP-based environments, and front-end PC-to-phone applications for IP-based interconnect partners. In many ways, the back-to-back GW area functions just like a normal non-SS7 POP, with the implications discussed in the subsections that follow.

Signaling

Back-to-back GWs simply need to be configured with similar TDM signaling types. Platforms considered for back-to-back configurations are the Cisco 3600 series, Cisco AS5300, Cisco AS5350, and Cisco AS5400.

Voice Quality/Bearer Issues

Voice quality suffers especially in the case of tandem compression. The addition of back-to-back GWs introduces additional postdial delay as well as added latency for all calls. There is even greater impact if more than one back-to-back GW zone is traversed. Fax relay may also suffer. Modem passthrough is highly unreliable, and as a result is not supported in scenarios that employ back-to-back GWs.

Note Where two back-to-back GWs are used, six routers are used in the path of a VoIP call(including the ingress and egress GWs at the network edge). Consequently, the end-to-end delay can approach 250 ms, with undesirable effects on QoS. In addition, in fax calls the tones transmitted across two back-to-back GWs are not received correctly at the terminating GW, as three encode/decode sequences are involved. Consequently, Cisco recommends that no more than one back-to-back GW be used in an end-to-end call.

Dial Plan

The back-to-back GW is responsible for manipulating digits and tech prefixes to fit into the general GK and DGK dial plan, as discussed in Basic Dial Plan, page 4-7. It also includes separating ingress and egress GWs in the GK call routing table. The extent of these considerations depends upon the application and the DGK/GK dial plan design. Dial plan responsibilities are discussed in more detail in the specific application templates presented later in this document.

Billing

One of the main benefits of the back-to-back GW is that it establishes a point in the call-signaling path from which to bill for IP-to-IP call topologies. The back-to-back GW largely functions just as a normal POP GW. Billing options vary by application type, and are discussed in more detail in the specific application templates presented later in this document.

Fault Tolerance

If the wholesaler desires, the back-to-back GWs may be configured to support an alternate GK with which it will register should the primary GK fail, just as with a normal TDM POP GW. For more information about fault tolerance and its requirements across functional areas, see Fault Tolerance, page 4-8.

Security

Back-to-back GWs have the same security options and implications as do normal POP GWs. For more information about security token mechanisms across functional areas, see Security, page 4-9.

Network Management

Back-to-back GWs have the same responsibilities as in a normal TDM POP.

TFTP Server Access

This is the same for back-to-back GWs as in a normal TDM POP.

Call Topologies and SP Interconnection Methods

Wholesale providers need to interconnect with other SPs, ITXCs, and other entities in order to offer minutes aggregation and termination, ASP termination, or card services. The interconnection method used is referred to as a *call topology.* A key question must be answered before a Cisco Wholesale Voice Solution can be initiated:

> *What is the interconnection type between the wholesale provider and other service providers?*

The appropriate template that the SP will use to deploy the Cisco Wholesale Voice Solution is determined by the functional areas required, as previously discussed, as well as the call topology and SP interconnection methods that are used.

Figure 1-1 showed the possible interconnection options that the wholesale VoIP service provider (SP) may need to accommodate at both the call ingress and egress sides. The possibilities are TDM at both ends, IP at both ends, and TDM at one end and IP at the other. The answer to the above question is determined by the nature of the demarcation between the wholesaler and those partner providers.

The call topology influences the configuration needs of the other functional groups within the network in supporting a given application. For example, if the wholesale provider enables simple carrier interconnect between an ASP and an IXC, then an IP-to-TDM topology would be used. The wholesale provider would then have to address the configuration considerations for that application, such as call routing, as well as shared support services for billing, settlement, and security in a way that is particular to that topology type.

Carrier Interconnect Scenarios

Consider the following simple carrier-interconnect scenarios. Only raw telephony minutes are exchanged, without the need for prompts, messages, or other user interaction. However, attention must be paid to routing, billing, and settlements.

1. Wholesaler sends traffic to an IXC, LEC, or other wholesaler over TDM.

   This is fairly straightforward. The wholesaler must ensure that the POP supports the offered TDM signaling type. Billing may be done at the terminating GW through AAA RADIUS, or on the TDM switch that provides the interconnect.

2. Wholesaler receives traffic from an IXC, LEC, or other wholesaler over TDM.

   Receiving traffic is fairly straightforward, with billing essentially as above. However, the wholesaler will most likely need to identify the source carrier in billing records. To determine which TDM interconnect partner originated the call, source-carrier information can be collected on either the originating GW or the TDM switch.

Note The Cisco Wholesale Voice Solution does not support propagating source identity information to the terminating GW, such as for use in CDRs or in an ingress customer service record (CSR) application on the terminating GW.

3. Wholesaler sends traffic to an ITSP or ASP over IP.

   IP interconnections complicate configurations. The wholesaler must choose how to route calls over IP to the terminating ITSP, ensure interoperability in bearer transport (by supporting, voice, fax, and modem calls), settle calls properly (through either OSP or CDR mediation), and provide support for desired security options (for example, through either H.235 or OSP).

4. Wholesaler receives traffic from an ITSP or ASP over IP.

   In addition to the considerations in the preceding scenario, the wholesaler must identify the source carrier for billing purposes. This is not trivial for a GK-based interconnection, as the wholesaler does not own the CDRs of the originating GW at the ITSP/ASP. The wholesaler must therefore extract that information at the terminating GW under that GW's administrative control (by means of either OSP or AAA).

   Alternatively, in interconnects that use OSP, the wholesaler may either own an OSP server or depend on a third-party clearinghouse OSP server to provide information about the originating carrier for settlement.

5. Wholesaler provides IP-to-IP transit network between two ITSPs/ASPs.

   This is a special case in which the wholesaler owns neither the originating nor the terminating GW, but rather simply provides transit between two ITSPs/ASPs. A wholesaler providing this service must therefore insert itself into the call signaling path to bill and settle calls, as well as to obscure originating carrier information from the terminating carrier and vice-versa. The wholesaler must also be aware of potentially different bearer transport options (such as codec types) and security options between the two interconnecting networks for which the wholesaler is providing transit.

   To accommodate this application, the Cisco Wholesale Voice Solution applies the concept of a back-to-back GW functional area.

The above scenarios, and others, can be accommodated by four fundamental call topologies:

- Topology 1: Originating TDM/Terminating TDM
- Topology 2: Originating TDM/Terminating IP

- Topology 3: Originating IP/Terminating TDM
- Topology 4: Originating IP/Terminating IP (Transit VoIP Network)

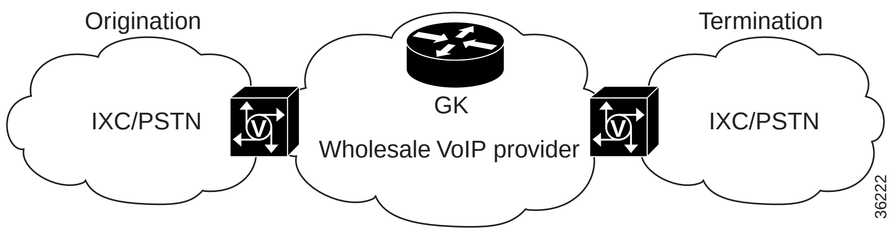
These topologies, which are the foundations of a solution design, are discussed below. Some application examples are included in each description.

Topology 1: Originating TDM/Terminating TDM

Description

This is a single administrative domain, and is the most fundamental call topology. Here the wholesale service provider both receives and terminates traffic between other service providers (for example, IXCs or LECs) over TDM interfaces at each end. Figure 2-6 illustrates this topology.

Figure 2-6 Originating TDM/Terminating TDM



Considerations

Because the interconnect is confined to TDM interfaces on GWs that are administered by the wholesale provider, deployment considerations with respect to routing, security, billing, and settlement are fairly straightforward (assuming no CSR applications are required).

The provider must take the following issues into account:

- The wholesaler must ensure that the POP supports the offered TDM signaling types.
- Transparency of bearer traffic (such as voice, fax, or modem passthrough) must be maintained

Applications

Applications for Topology 1 include the following:

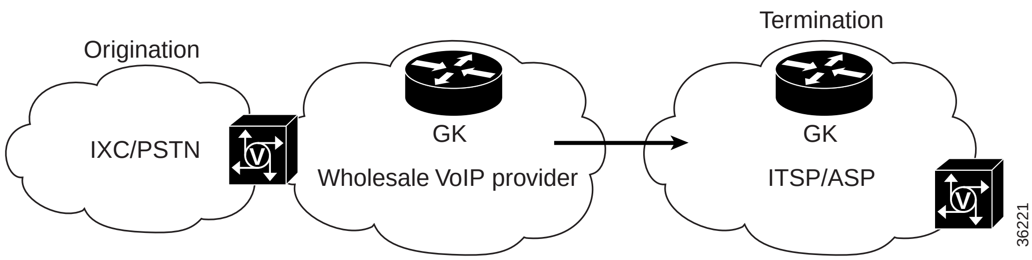
- Card services
- IXC-to-IXC interconnect
- IXC offload
- LEC-to-LEC interconnect (simple toll bypass)
- LEC-to-IXC interconnect
- Limited RADIUS-based CSR

Topology 2: Originating TDM/Terminating IP

Description

The wholesaler may want to increase call volume or coverage by adding interconnections with other IP-based service providers. Here the wholesale service provider receives traffic from other service providers (such as IXCs or LECs) over TDM interfaces. Figure 2-7 illustrates this topology.

Figure 2-7 Originating TDM/Terminating IP



Considerations

If the provider cannot terminate the call within its own network POPs, it can send traffic to other service providers (such as ITSPs or ASPs) over IP.

In addition to the TDM-related issues discussed under Topology 1, the wholesale provider must take into account the following issues related to the IP interconnect:

- Call routing
- Interoperable bearer transport (such as codec types—are they compatible end to end?)
- Billing
- Settlement
- Security

Applications

Applications for Topology 2 include the following:

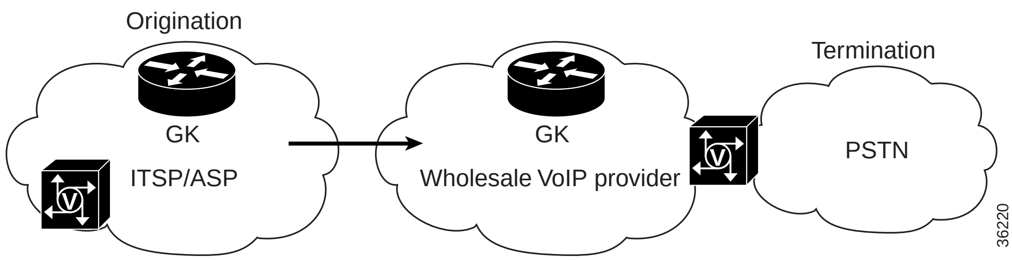
- Card services
- LEC-to-ASP interconnect
- LEC-to-ITSP interconnect (simple toll bypass)
- IXC-to-ASP interconnect
- IXC-to-ITSP interconnect

Topology 3: Originating IP/Terminating TDM

Description

This is essentially the same as Topology 2, except that the call direction is reversed. Here the wholesale service provider receives traffic from other service providers (such as ITSPs or ASPs) over IP and terminates traffic at its POPs to other service providers (such as IXCs or LECs) over TDM interfaces. Figure 2-8 illustrates this topology.

Figure 2-8 Originating IP/Terminating TDM



Considerations

Because the wholesale provider is now receiving traffic from other providers over an IP interconnect, the wholesale provider must take into account the following issues:

- Call routing
- Originating carrier identification (for billing and settlement)
- Interoperable bearer transport (such as codec types—are they compatible end to end?)
- Security

Applications

Applications for Topology 3 include the following:

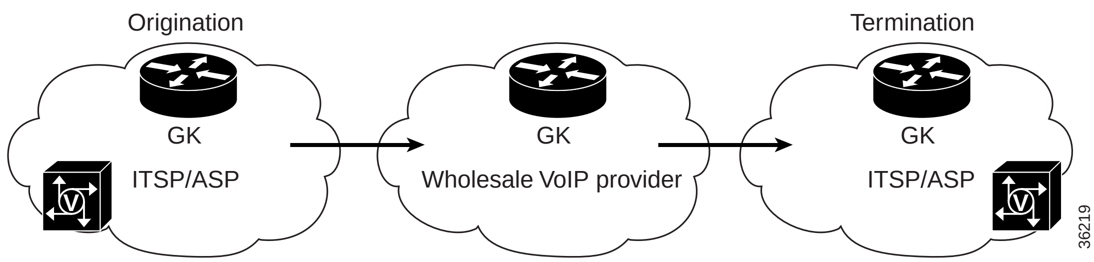
- ITSP-to-LEC interconnect (toll bypass)
- ASP-to-LEC interconnect (toll bypass)
- ITSP-to-IXC interconnect
- ASP-to-IXC interconnect

Topology 4: Originating IP/Terminating IP (Transit VoIP Network)

Description

Finally, the wholesaler service provider may want to provide transit between different IP-based interconnect partners. Here the wholesaler exchanges traffic among other service providers (such as ITSPs or ASPs) over IP connections only. Typically, the wholesaler receives traffic from an ITSP or ASP. If the wholesaler cannot terminate the call at one of its own POPs, it passes the call off to another service provider. (This is simply a transit VoIP network.) Figure 2-9 illustrates this topology.

Figure 2-9 Originating IP/Terminating IP (Transit VoIP Network)



Considerations

In sending and receiving traffic between two IP interconnects, the wholesale service provider has increased challenges with respect to the following issues:

- Call routing
- Carrier identification
- Billing
- Settlement
- Security
- Potentially “masking” (obscuring) originating-carrier information from the terminating carrier

Applications

Applications for Topology 4 include the following:

- ASP-to-ITSP interconnect
- ASP-to-ASP interconnect
- ITSP-to-ITSP interconnect
- ITSP-to-ASP interconnect

Deployment Templates

Finally, in addition to the above interconnection possibilities, deployment templates are determined on the basis of whether a DGK or OSP server is used. The following scenarios are considered:

1. TDM to TDM
2. TDM to IP with DGK-based interconnect
3. TDM to IP with OSP-based interconnect
4. IP to IP with DGK-based interconnect
5. IP to IP with OSP-based interconnect

For the two service types, these scenarios are discussed in detail, in the above sequence, in the following chapters:

- Chapter 5, "Service A: Minutes Aggregation and Resale (Including ASP Termination)"
- Chapter 6, "Service B: Card Services (Prepaid and Postpaid)"

CHAPTER 3

Solution Components

To support the various topologies and applications of the Cisco Wholesale Voice Solution, both Cisco and third-party components are required. Some components are required in all topologies. Other components, although optional, are mandatory for certain solutions (such as those requiring SS7 signaling). This chapter lists and discusses the actual components, both Cisco's and those of third parties, that can be used to implement a Cisco Wholesale Voice Solution.

This chapter presents the following major topics:

- Major Components
- Additional Components (for Shared Services)

Major Components

Table 3-1 summarizes the major components, all of which are Cisco components. Required components are listed first, followed by optional components.

Note Refer to the most current *Cisco Wholesale Voice Solution Release Notes* for the latest information regarding IOS versions and the platforms they run on.

Table 3-1 Major Components of the Cisco Wholesale Voice Solution

Component	Description
Cisco Voice Gateways (GWs)	*Required*. Support small- to large-scale interconnects with wholesaler's TDM-based customers. Platforms include the Cisco 3600 series, Cisco AS5300, Cisco AS5350, and Cisco AS5400, as well as supporting network modules. The Cisco 3600 series is not used in applications requiring SS7 signaling. For more information see Cisco Voice Gateways, page 3-2.
Cisco H.323 Gatekeepers (GKs)	*Required*. Allow network to be scaled to large sizes. Platforms include the Cisco 3660 and Cisco 7200. For more information see Cisco H.323 Gatekeepers and Directory Gatekeepers, page 3-4.
Cisco H.323 Directory Gatekeepers (DGKs)	*Required*. Supplement GKs in allowing network to be scaled to large sizes. Platforms include the Cisco 3660 and Cisco 7200. For more information see Cisco H.323 Gatekeepers and Directory Gatekeepers, page 3-4.
Cisco Signaling Controllers	*Optional*. However, these components are *required* in SS7 interconnect solutions. The supported platform is the Cisco SC2200 running Release 7.4 (Cisco SS7 Interconnect for Voice Gateways). For more information see Cisco Signaling Controllers, page 3-5.
Cisco SS7 Signaling Link Termination (SLT) Systems	*Optional*. However, these components are *required* in SS7 interconnect solutions. The platform is the Cisco 2611. For more information see Cisco SS7 Signaling Link Termination Systems, page 3-5.

The above components are discussed in the sections that follow.

Cisco Voice Gateways

The Cisco Wholesale Voice Solution requires a range of small-, medium-, and large-scale PSTN interconnects with the wholesaler's TDM-based customers (typically IXCs, PTTs, or other wholesalers), depending on anticipated call volumes. Platforms include the Cisco 3600 series, the Cisco AS5300, the

Cisco AS5350, and the Cisco AS5400, along with various supporting network modules. To offload traffic, similar interconnects may be required. The GWs may handle their own signaling, or they may provide IMTs and receive external SS7 signaling through a Cisco SC2200 running Cisco SS7 Interconnect for Voice Gateways, with Q.931 signaling backhaul.

Note The Cisco 3600 series is not used in applications requiring SS7 signaling.

For the modules and signaling types supported by the Cisco VoIP GWs, see Table 3-2.

Table 3-2 Cisco VoIP Gateways, Interface Modules, and Supported Signaling Types

Platform	Interface Modules	Signaling Types
Cisco 3600 series GWs (Cisco 3620, Cisco 3640, Cisco 3660)	NM-2V • Analog FXO WIC • Analog E&M WIC • BRI WIC	Analog FXO ground start, loop start, immediate start Analog E&M types I through V BRI (ETSI, NI-2)
	NM-HDV-1 T1/E1 NM-HDV-2 T1/E1 T1/E1 WVIC	PRI (ETSI, NI-2) T1 CAS (FGB[1], FGD) E1/R2 E1/R2—MF
Cisco AS5300, Cisco AS5350	Quad T1 Quad E1	PRI (ETSI, NI-2) T1 CAS—(FGB, FGD) E1/R2 E1/R2—MF Cisco SS7 Interconnect for Voice Gateways (Q.767, TR-113, China TUP)
Cisco AS5400	Octal T1/E1	CT1/CE1/PRI (ETSI, NI-2) T1 CAS (FGB, FGD) E1/R2 E1/R2—MF DS3 Cisco SS7 Interconnect for Voice Gateways (Q.767, TR-113, China TUP)

1. FG = feature group

Note The Cisco Wholesale Voice Solution does not support GW platforms that use MGCP call signaling. In addition, only Cisco SS7 Interconnect for Voice Gateways is used.

Cisco H.323 Gatekeepers and Directory Gatekeepers

Cisco H.323 GKs and DGKs constitute the gatekeeper core. GKs are mandatory network elements used to scale a wholesale network to large sizes. Directory GKs (DGKs) further supplement network scalability and are mandatory if GK-based carrier interconnect is desired. The Cisco GKs and DGKs supported by the Cisco Wholesale Voice Solution are essentially IOS software images loaded onto a dedicated Cisco 3660 or Cisco 7200 series machine.

Note The Cisco Wholesale Voice Solution does not support integrated GKs and GWs.

Cisco GKs perform the following tasks.

Resource Management

Cisco GKs determine the health of H.323 gateways by monitoring registration and unregistration (RRQ/URQ) messages and resource availability indicators (RAIs).

Call Routing

Cisco GKs provide call routing based on destination E.164 addresses. They may use their knowledge of local GW health levels to make routing decisions in order to increase the network availability of the GWs. Cisco GKs may also route calls between remote GKs within the same administrative domain by means of inter-GK LRQ RAS messages. Similarly, Cisco DGKs may also route calls to other carrier administrative domains by means of LRQ RAS messages.

Security

Cisco GKs, in conjunction with an external server (such as RADIUS), may be used for secure call admission in intradomain call scenarios (calls within the same service provider's domain). Cisco GKs also have limited applications in implementing interdomain security functions for calls sent between carriers by means of IP interconnect.

External GKTMP Applications

Cisco GKs may act as a control point from which an application server can affect call routing, number translation, call admission/blocking, and the like. These application servers interface with a Cisco GK or DGK by means of GKTMP.

Note The Cisco Wholesale Voice Solution does not specify any particular GKTMP application, but does not exclude any from being used.

CDR Generation

Cisco GKs have limited abilities to generate CDR reports for calls either in addition to or instead of generating them from the GW. This is an option if the wholesaler either does not own the GWs at a POP, or simply wants to reduce the amount of messaging overhead associated with AAA in its smaller POPs. However, billing in this manner has limitations.

Cisco Signaling Controllers

These optional components are used in SS7 interconnect solutions. The supported platform is the Cisco SC2200 running Cisco SS7 Interconnect for Voice Gateways.

Cisco SS7 Signaling Link Termination Systems

These optional components are used in SS7 interconnect solutions. Signaling link termination (SLT) systems are Cisco 2611 platforms that are capable of terminating MTP1 and MTP2 SS7 layers and backhauling MTP3 and higher SS7 layers to the Cisco SC2200 in an SS7 interconnect solution.

Additional Components (for Shared Services)

This section discusses the additional components that support shared services. Table 3-3 summarizes these components, which are provided by third parties.

Table 3-3 Additional Components of the Cisco Wholesale Voice Solution

Component	Description
RADIUS/OSS Servers	*Required*. Ecosystem Partner OSS servers interface with Cisco GW and GK components through AAA RADIUS. For more information about these components, see RADIUS/OSS Servers, page 3-6.
Ecosystem Partner H.323 GKs[1]	*Optional*. These GKs are used on the edge of the network to complement the Cisco GK/DGK infrastructure and host a variety of applications. Applications will vary among Ecosystem Partners. For more information about these components, see Ecosystem Partner H.323 GKs, page 3-7.
Gatekeeper Application Servers[2]	*Optional*. These host enhanced call-routing applications and interface with the Cisco wholesale VoIP network through Cisco GKs or DGKs, by means of the GKTMP interface specification. For more information about these components, see Gatekeeper Application Servers, page 3-8.
OSP Servers	*Optional*. These are used in simple carrier interconnect scenarios. The use of OSP for secure settlement transactions requires a clearinghouse entity, or at least a dominant carrier in the interconnect relationship that administers the OSP server. For more information about these components, see OSP Servers, page 3-8.

Table 3-3 Additional Components of the Cisco Wholesale Voice Solution (continued)

Component	Description
Prompt Servers	*Optional*. These maintain a prompt database for GWs running interactive voice response (IVR) functionality for applications such as card services. For more information about these components, see Prompt Servers, page 3-9.
TFTP Servers	*Optional*. These store a variety files that do not need to reside on the local machine. For more information about these components, see TFTP Servers, page 3-9.
Management Systems	*Optional*. Both Cisco and third-party management systems are supported. These include element management systems (EMSs) and network management systems (NMSs). For more information about these components, see Management Systems, page 3-9.

1. The Cisco Wholesale Voice Solution does not require or specify the use of these components, but the architecture does not exclude their use.
2. The Cisco Wholesale Voice Solution does not require or specify the use of specific GKTMP applications, but the architecture does not exclude their use.

The above components are discussed in detail in the sections that follow.

RADIUS/OSS Servers

OSS servers interface with Cisco GW and GK components via AAA RADIUS VSAs (vendor-specific attributes), and are *required* elements of the wholesale network. Current examples include Cisco Secure and billing platforms such as those provided by Cisco Ecosystem Partners MIND/CTI (http://www.mindcti.com) and Belle Systems (http://www.bellesystems.com).

Note In the Cisco Wholesale Voice Solution, Cisco Secure does not support applications such as debit card, which are dependent on VSAs. For more information about RADIUS VSAs, see *RADIUS Vendor-Specific Attributes Voice Implementation Guide* at http://www.cisco.com/univercd/cc/td/doc/product/access/acs_serv/vapp_dev/vsaig3.htm.

VSAs can be used to achieve the following functions, as discussed below:

- CDR Collection and Billing System Front-Ending
- User Authentication and Authorization
- Application Hosting
- Security
- Settlement

CDR Collection and Billing System Front-Ending

Cisco GWs send call start/stop records to a RADIUS server by means of AAA. The billing application can extract these records to generate CDRs. CDRs may be shared between carriers as a method of settlement through billing-system mediation applications.

User Authentication and Authorization

For card services, an AAA RADIUS server may validate end users on the basis of ANI or username and password combinations. The AAA interaction occurs directly on the Cisco GW.

Application Hosting

A Cisco GW may run a call script that interacts with an application mounted on the RADIUS server. The server is capable of manipulating call information by means of VSAs in AAA. An example would be a debit card application. The AAA server interacts with a debit-card billing application to determine account balances, call rates, and the time remaining for an individual user. This information is sent to the GW script in AAA VSAs.

Note In the Cisco Wholesale Voice Solution, Cisco Secure does not support applications such as debit card, which are dependent upon VSAs.

Security

GKs can administer H.235 security options to perform secure endpoint registrations.

Note The Cisco Wholesale Voice Solution supports Cisco H.235 security for *registration only*. For reasons of AAA latency, H.235 security is not supported on a per-call basis. When per-call security is required, Cisco recommends that you use access lists.

Settlement

Some billing system vendors support interdomain settlement based on CDRs that are collected from each local domain. This offers a viable alternative to OSP in some cases. Mediation vendors such as XACCT (http://www.xaccct.com) also provide servers dedicated to settling CDRs between different vendors' billing systems. These are known as *mediation servers*, and are optional components in a wholesaler network.

Ecosystem Partner H.323 GKs

These GKs may be optionally used on the network fringe to compliment the Cisco GK and DGK infrastructure to host a variety of applications. Individual applications will vary among Ecosystem Partners.

Note Cisco has identified leading solution providers with which to partner in offering end-to-end benefits. This forms what is known as the Cisco New World Ecosystem. For more information about the Ecosystem Partner community, visit http://www.cisco.com/public/Partner_root.shtml.

The Cisco Wholesale Voice Solution neither requires nor specifies where and when these partners should be used on a wholesale network. However, the architecture does not exclude them from being inserted into the wholesaler's network.

Gatekeeper Application Servers

Enhanced call routing applications may optionally reside on an external server. They can interface with a Cisco wholesale VoIP network through Cisco GKs or DGKs that use the GKTMP interface specification.

Note The Cisco Wholesale Voice Solution does not specify any particular GKTMP applications, but the solution architecture does not exclude adding them into the wholesaler's network.

OSP Servers

To support carrier interconnect, wholesale providers may choose to use OSP servers. (For more information about OSP, see Appendix A, "Open Settlements Protocol (OSP) Clearinghouse Solution.") The use of OSP for secure settlement transactions involves a clearinghouse entity or at least a dominant carrier in the interconnect relationship that administers the OSP server.

Note TransNexus, Inc. (http://www.transnexus.com) currently provide OSP-based clearinghouse services.

OSP servers perform the following functions, as discussed below:

- Authentication of GWs or Carriers
- Call Authorization
- Call Routing
- CDR Collection
- CDR Correlation and Settlement

Authentication of GWs or Carriers

By using a secure exchange of certificates, an OSP server can validate whether or not an originating or terminating carrier's GW is a valid participant in the OSP interconnect.

Call Authorization

An OSP server generates an access token for each call sent from an originating GW into the OSP-based interconnect. The originating GW includes this token in the SETUP message to the terminating GW. Upon receiving the SETUP, the terminating GW may then send the token back to the OSP server for validation, or may perform the validation locally.

Call Routing

The OSP server provides the originating GW with a terminating GW selected from among registered OSP endpoints.

CDR Collection

After a call has ended, OSP usage indicators are sent to the OSP server from both the originating and terminating endpoint. The OSP server uses this information to generate a CDR.

CDR Correlation and Settlement

Once CDRs are collected, the OSP server may interface with a billing application to generate settlement billing between the two interconnecting carriers.

Prompt Servers

A prompt server is an optional component that maintains a prompt database for GWs running IVR functionality for applications such as card services. If they are not too big, prompt databases may be stored locally on the GW in flash memory. Larger prompt databases, such as those needed when there are many branded retailers or when many languages must be supported, may be downloaded dynamically as needed from a TFTP prompt server. TFTP servers are generic third-party devices that can be hosted on a wide variety of platforms.

TFTP Servers

TFTP servers are used to store a variety of files that do not need to reside on a local machine, and that would otherwise take up available limited memory on that machine. These files can be downloaded as needed. Example files include audio (IVR) files, IOS files, and configuration files (including dial peers).

Management Systems

The Cisco Wholesale Voice Solution supports both element and network management systems.

Element Management Systems

Element management systems (EMSs) are optional components that are used for managing or provisioning other components in the solution. CVM 2.0.2 provides limited provisioning support and is currently the only EMS supported by the Cisco Wholesale Voice Solution.

Network Management Systems

Network management systems (NMSs) are optional components that are used for network monitoring, fault management, trap correlation, and reporting. Any NMS can extract this information from wholesale components by means of SNMP. The Cisco Wholesale Voice Solution recognizes CiscoWorks Internet Protocol Manager (IPM) to monitor network QoS, and Cisco Info Center (CIC) to provide fault management and trap correlation. For reporting, it is possible for third-party vendors of reporting applications to provide reports by interfacing with CVM 2.0.2. The Cisco Wholesale Voice Solution recognizes Trinagy (http://www.trinagy.com) as one of these vendors.

Designing a Solution

You have been introduced to the many variables that determine how to configure and implement a Cisco Wholesale Voice Solution. This chapter shows you how to determine the appropriate solution for a given set of circumstances.

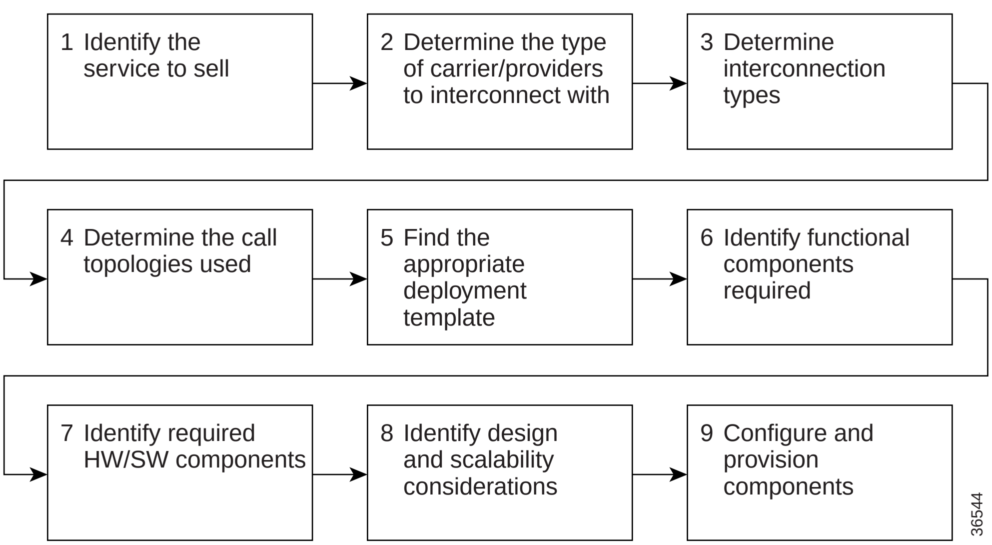
This chapter presents the following major topics:

- First Steps
- Design Issues and Choices
- Service Designs

First Steps

Begin by asking the right questions and identifying key issues. Figure 4-1 illustrates the decision blocks necessary to determine a particular Cisco Wholesale Voice Solution.

Figure 4-1 Steps in Designing a Solution



These steps are discussed below.

Step 1 Identify the service to sell.
(See Services Supported, page 1-4.)

- Service A: Minutes Aggregation and Resale (Including ASP Termination)?
- Service B: Calling Card Service (Prepaid and Postpaid)?

Step 2 Determine the type of carriers/providers to interconnect with.
(See Call Topologies and SP Interconnection Methods, page 2-11.)

- ITSP?
- ASP?
- IXC (PSTN)?

Step 3 Determine the interconnection types.
(See Call Topologies and SP Interconnection Methods, page 2-11, and Design Issues and Choices, page 4-3.)

- TDM?
- IP?

Step 4 Determine the call topologies used.
(See Call Topologies and SP Interconnection Methods, page 2-11.)

- Topology 1: Originating TDM/Terminating TDM?
- Topology 2: Originating TDM/Terminating IP?
- Topology 3: Originating IP/Terminating TDM?
- Topology 4: Originating IP/Terminating IP (Transit VoIP Network)?

Step 5 Find the appropriate deployment template.
(See Deployment Templates, page 2-17, and Design Issues and Choices, page 4-3.)

Note OSP or DGK peering may be required. Refer to IP Interconnect Variations, page 4-3, for more information.

- TDM to TDM?
- TDM to IP?
- TDM to IP with OSP?
- IP to IP with DGK?
- IP to IP with OSP?

Step 6 Identify the functional components required.
(See Functional Areas, page 2-1, and Design Issues and Choices, page 4-3.)

- Non-SS7 POP?
- SS7 POP?
- Back-to-Back GW Area?
- Gatekeeper Core?
- Shared Support Services?

Step 7 Identify the required hardware/software components.
(See Chapter 3, “Solution Components.”)

Step 8 Identify design and scalability considerations.
(See to Design Issues and Choices, page 4-3.)

Step 9 Configure and provision components.
(Refer to the *Cisco Wholesale Voice Solution Design and Implementation Guide*.)

Design Issues and Choices

This section discusses in detail a variety of issues choices that will vary with each application.

Because of the many ways in which multifunctional groups can interact, a variety of issues deserve discussion in greater detail. Among these are the following:

- IP Interconnect Variations
- Call Routing and Billing/Settlement
- Basic Dial Plan
- Network Management

IP Interconnect Variations

In addition to the four fundamental topologies discussed in Call Topologies and SP Interconnection Methods, page 2-11, wholesalers can interconnect with other IP-based service providers (ITSPs and ASPs) in one of two ways:

- DGK-Based Interconnect
- OSP-Based Interconnect

Each method has its own provisioning requirements.

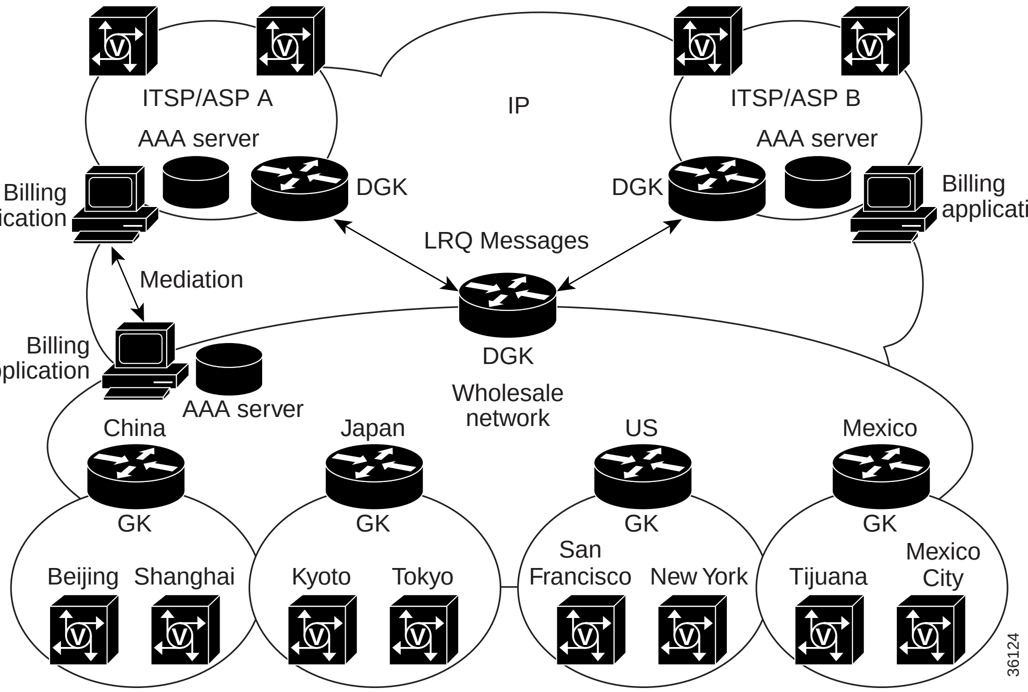
DGK-Based Interconnect

In this method, the wholesaler provisions call routing between its IP interconnect partners by peering DGKs to which it sends LRQ RAS messages. The wholesaler may direct certain destination patterns to specific interconnect partners. These destination patterns could have potentially been modified upon ingress into the wholesaler network to provide limited ingress CSR applications. Additionally, the wholesaler may also use sequential LRQ features to provide limited egress CSR applications.

With DGK-based interconnect, the wholesaler benefits from centralizing route provisioning in the DGK, rather than pushing it to the edge GWs as with OSP. However, billing/settlement functions and security options are processes external to call routing that require some configuration in the GWs, GKs, and related shared-services components.

For a large service provider with many POP GWs, provisioning complexities may decide that this is the more attractive option for interconnect. Figure 4-2 illustrates a DGK-based interconnect with other ITSP/ASP partners.

Figure 4-2 DGK-Based Interconnect with Other Service Partners



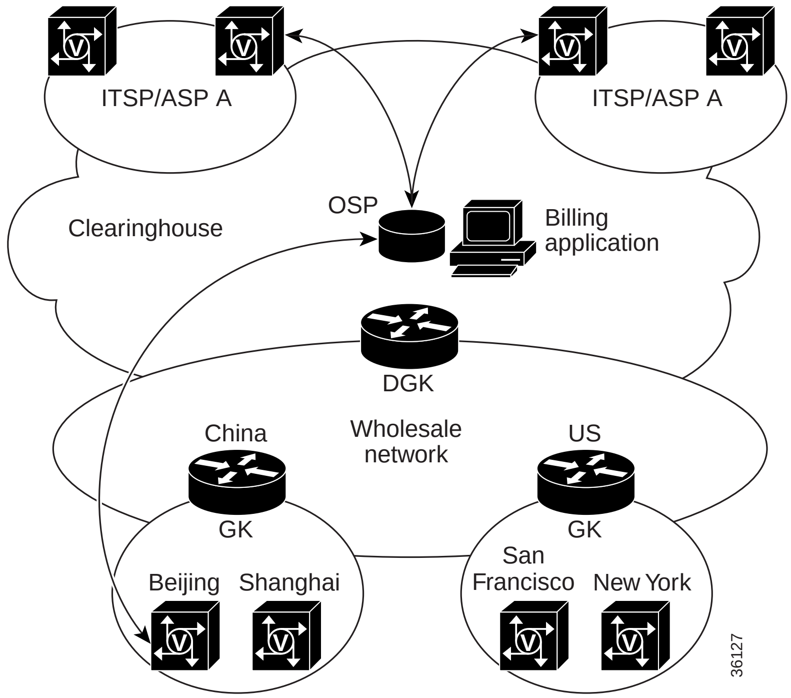
OSP-Based Interconnect

In this method, an OSP server performs call routing, billing/settlement, and security functions. However, OSP-based interconnect requires additional provisioning. All edge GWs must be registered with the OSP server, and rotary dial-peer failover must be provisioned to route calls through the OSP interconnect.

OSP may be an attractive interconnect option for a wholesaler who wants to combine call routing, security, and billing/settlement into one architecture. However, in current Cisco implementations, limitations with OSP deployments require extensive provisioning in the GWs, so that they can interact with the required shared services, support the dial plan architecture, and cover termination caveats.

Figure 4-3 illustrates OSP-based interconnect with other ITSP/ASP service partners.

Figure 4-3 OSP-Based Interconnect with Other Service Partners



Call Routing and Billing/Settlement

Originating and terminating providers may connect to the wholesaler through either TDM or IP. The call routing and billing/settlement functions determine whether the wholesaler will be DGK or OSP based.

Wholesale providers have the option to interconnect routes with other service providers using a DGK. In this model, the wholesaler and the interconnect partner configure the DGKs to exchange LRQ RAS messages between their administrative domains, in order to resolve call routing for desired prefixes.

Call Routing

The DGK maintains an overall routing table of destination patterns and the corresponding local GKs that support them. The DGK simply forwards LRQ requests to the local GK that handles that destination pattern.

This use of GKs and DGKs allows the addition of new GK zones, POPs, and certain types of IP interconnect partners with minimal impact on dial plan provisioning. Changes are isolated to the local GK and the DGK. The rest of the elements in the network are untouched. Often, the level of dial plan resolution at the DGK can be simplified. For example, a DGK may know to route all calls beginning with a country code of 1 to the local U.S. GK. The local US GK can then expand selection to more digits (such as NPA or NPA-NXX) to route the call to the proper terminating gateway.

There are two design choices for call routing between IP service providers:

- DGK-Based Call Routing
- OSP-Based Call Routing

DGK-Based Call Routing

DGK-based call routing uses LRQ (Location Requests) to request the terminating GW IP addresses. An LRQ is sent from the originating SP's DGK to the terminating service provider's DGK. SPs would use the DGK method of call routing where the originating and terminating service providers are trusted peers.

OSP-Based Call Routing

OSP-based call routing uses a separate entity called an *OSP clearinghouse*, which maintains OSP servers. These OSP servers contain the prefix call routing tables of all SPs that subscribe to the clearinghouse. The originating GW, in this case, sends an ARQ (Authorization Request, an OSP message) to the OSP server, and the OSP server responds with an ARP (Authorization Response) containing a list of possible IP addresses of the terminating GW plus a security token. This token is included in the SETUP message to provide security validation at the terminating GW. The SP would use the OSP method of call routing when carriers want a third party to provide the billing and settlement.

Billing/Settlement

To account for service properly, the wholesaler must accurately identify the originating carrier and terminating carrier for calls. The degree of difficulty for this varies depending upon the call topology used. Furthermore, the usage indicators must be extracted from a reliable source. This implies that the devices supplying call usage indicators are somewhere within the H.225 call-signaling path. In the Cisco Wholesale Voice Solution, if billing is desired, this means the wholesaler must own at least one GW in any given conversation.

There are two design options for billing and settlement:

- AAA RADIUS-Based Billing/Settlement
- OSP-Based Billing/Settlement

These options may be used either individually or in conjunction with each other and will directly depend upon the method of interconnect. Though differing in protocol, each method addresses the same basic needs for call accounting.

AAA RADIUS-Based Billing/Settlement

The wholesaler implements AAA billing for any intradomain traffic because OSP is designed to bill for interdomain calls only. AAA may be also be used for interdomain calls if interconnect is not handled by an OSP server, but rather by a peering DGK relationship. In this model, the billing application correlates the usage records to generate CDRs. The wholesaler then distributes these CDRs to customers in the form of a periodic bill. Customers can verify this bill against their own records before ultimately exchanging money and settling the bill. There are various mediation vendors that help automate the verification and settlement stages, but the Cisco Wholesale Voice Solution does not specify any. Details regarding from which GWs AAA accounting records are collected and how they are correlated are addressed in the specific deployment templates presented in the chapters that follow.

OSP-Based Billing/Settlement

For interconnects using OSP, the wholesaler may either own an OSP server or depend upon a third-party OSP clearinghouse server to provide accounting services. The OSP server receives accounting information from the wholesaler GW much as in the AAA RADIUS case. Since the OSP server receives usage indicators from both GWs across administrative domains, the OSP server gets accurate terminating and originating carrier information. The OSP server may then correlate the usage records to generate CDRs. These CDRs may be distributed as periodic bills to customers. Customers can verify their bills against their own records before ultimately exchanging money or settling the call. To provide personal accounting records for verification, the wholesaler may use parallel AAA accounting records.

In the case of an OSP clearinghouse, the clearinghouse may also provide a single interface for settlement between multiple providers, but at an added cost. Details regarding OSP billing mechanisms are provided in the specific deployment templates discussed in the chapters that follow.

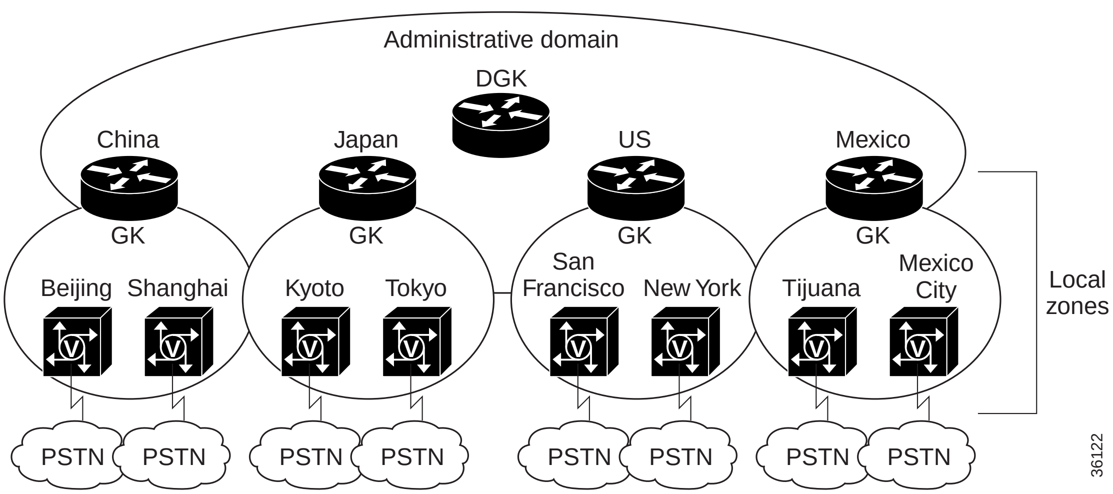
It is a possibility that a third party and not the wholesaler manages an interconnecting TDM POP. In this case, the wholesaler cannot depend upon GWs to send them CDR info. A wholesaler may therefore choose to do billing from either *only* the terminating GWs (if the wholesaler owns them) or *only* from the GK.

Billing from the GK has limitations. Cisco GKs can send call start/stop records to an AAA RADIUS server based upon receipt of ARQ and DRQ RAS messages from GWs. However, RAS messages are sent over UDP and not guaranteed to arrive at the GK. Furthermore, this method of billing lacks answer supervision. Therefore, billing is most reliable and accurate if performed at the gateway. The Cisco Wholesale Voice Solution recommends that GW-based billing be done whenever possible.

Basic Dial Plan

Within the wholesaler network, dial plan responsibilities are distributed among GWs, GKs, and DGKs. Because deployments of Cisco SS7 Interconnect for Voice Gateways leverage NI-2 type Q.931 backhaul signaling, the basic H.323 dial plan architecture is the same regardless of whether the POPs in the network are SS7 based, non-SS7 based, or a mixture of both. Figure 4-4 depicts a typical large-scale H.323 network design (the same figure as seen in Figure 2-2).

Figure 4-4 Typical Large-Scale H.323 Network Design



GWs deal with the local POP portion of the dial plan. This encompasses any digit manipulation needed to normalize numbers or implement local PSTN access rules. It also includes notifying a GK when resource availability indicator (RAI) thresholds are crossed in order to increase call completion rates. Furthermore, the GW may implement rotary dial peers to handle call-failover routing options, such as trying OSP if normal GK RAS call routing offers no possible termination.

For example, the provider may want the GW to notify the GK when its resource limits are nearly exhausted, thereby prompting the GK to select a different GW. Additionally, the provider may want to normalize numbers into a standard format before sending calls into the VoIP network (such as country code + area code + local number) to simplify route provisioning in the GKs and DGKs. Likewise, the provider may need to prefix/strip digits, as PSTN access rules require (such as prefixing any needed area or access codes), before sending calls out the TDM interfaces.

Local GKs monitor gateway health levels and maintain detailed routing tables, mapping destination patterns to specific terminating GWs within local zones. The local GKs can use features such as lightweight registration, RAI, and static gateway-priority assignments to influence gateway selection. For all other nonlocally supported destination patterns, the local GK configures a wild-card route to the DGK.

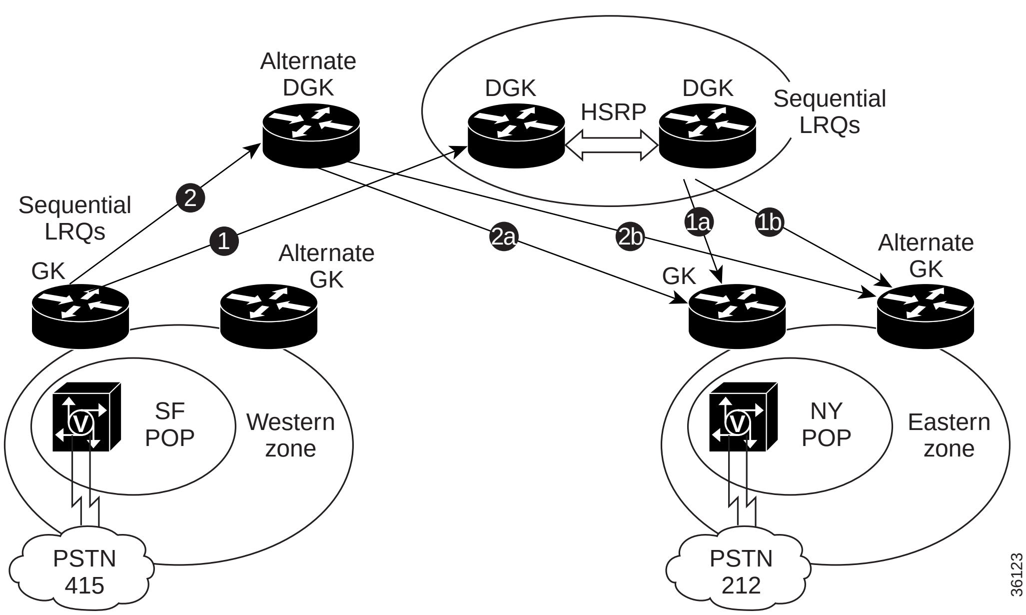
The DGK maintains an overall routing table of destination patterns and the corresponding local GKs that support them. The DGK simply forwards LRQ requests to the local GK that handles that destination pattern.

This use of GKs and DGKs allows the addition of new GK zones, POPs, and certain types of IP interconnect partners with minimal impact on dial plan provisioning. Changes are isolated to the local GK and the DGK. The rest of the elements in the network are untouched. Often, the level of dial plan resolution at the DGK level can be simplified. For example, a DGK may know to route all calls beginning with a country code of 1 to the local US GK. The local US GK can then expand selection to more digits (such as NPA or NPA-NXX) to route the call to the proper terminating GW.

Fault Tolerance

For intradomain calls and DGK-based IP interconnects, a wholesaler has the option of overlaying fault tolerance on to the basic H.323 VoIP network dial plan design. This is done by using a combination of Cisco IOS software features such as alternate GKs on the GW, HSRP on the DGK, and sequential LRQs on the GKs and DGKs. Figure 4-5 illustrates a fault-tolerant architecture using alternate GKs.

Figure 4-5 Fault-Tolerant Architecture using Alternate GKs



GWs may be configured to register to a primary GK and an alternate GK in the event that the primary GK fails. This implies that at any given time, a GW may be registered to either its primary or alternate GK. Since Cisco GKs do not communicate registration states between each other, sequential LRQs must be configured on the GKs and DGKs to accommodate zone fragmentation.

For example, a GK in the Western zone supports GWs in San Jose (408) and San Francisco (415). Under normal circumstances, when San Jose calls San Francisco, the route is resolved in the local primary GK. However, assume that San Jose fails over to the alternate GK while San Francisco remains on the primary GK. To continue to support regional call completion within the Western zone, the primary and alternate GKs must be provisioned to send local prefixes to each other if no local resources exist (for example, the terminating GW has failed over to the other GK). In this case, in order for San Francisco to complete calls to San Jose, the primary GK must know to send an LRQ for the San Jose prefix to the alternate GK. Similar provisioning is required on both primary and alternate GKs to support calls in both directions.

Provisioning is also required on the DGK to prevent zone fragmentation issues when calls are originated from other zones. For example, if San Francisco sends a call to New York, the DGK does not know with which GK (primary or alternate) the NY GW is registered. The DGK must be provisioned to send sequential LRQ requests to both primary and alternate terminating local GKs for all Eastern zone supported prefixes (messages 1a and 1b, which are addressed to both DGKs in the HSRP area). Similar provisioning is required for the Western zone prefixes to support calls in the other direction.

HSRP is used to provide fault tolerance for the DGK. However, HSRP failover detection may take some time during which no calls will be processed. To cover this case, local GKs may be configured to point to more than one DGK (an ALTDGK) for its wild-card route using sequential LRQs.

For example, the GK may point to an HSRP DGK pair as its primary option (message 1). If no response is received because HSRP failover has not been detected yet, the GK may initiate another LRQ (message 2) to an ALTDGK after a configurable timeout (100 to 1000 ms) to try to complete the call. During this time, calls will still be completed, although with additional postdial delay. The ALTDGK is configured in exactly the same way as the primary DGK HSRP pair (messages 2a and 2b).

Security

A wholesaler may choose to implement various security methods over its H.323 VoIP network. The selected security mechanism has different provisioning needs within multiple functional areas. For intradomain calls, the wholesaler may use complex access lists.

GWs in the wholesaler network may be provisioned with complex access lists to accept calls from only known entities. This is obviously neither scalable nor friendly to network changes or elements that use DHCP (Dynamic Host Configuration Protocol). For interdomain calls, the wholesaler may choose to implement either complex access-lists or H.235 OSP access tokens, depending upon the interconnect method used (DGK vs. OSP).

Note The Cisco Wholesale Voice Solution supports Cisco H.235 security for *registration only*. For reasons of AAA latency, H.235 security is not supported on a per-call basis. When per-call security is required, Cisco recommends that you use access lists.

Network Management

The Cisco Wholesale Voice Solution supports different NMS components, each addressing a certain task, and that can be assembled to provide network management. The tasks, along with the components that address them, are considered below.

Configuration Management

This deals with installing, provisioning, and tracking the provisioning of all components in the network.

Element Management

This deals with real-time monitoring of network components, plus the ability to control them by bringing them up or down.

Fault Management

This is the ability to detect faults from individual components and correlate them into a coherent picture of an overall system problem.

Reporting

The wholesaler or wholesaler's partners may want to generate reports related to system configuration, performance, and faults. In some cases, not all aspects of a component can be managed with available products. It may be necessary to manage devices through the CLI, or by using SNMP systems to manipulate the MIB directly.

Service Designs

This following two chapters provide a detailed discussion of issues that must be considered for the services of the Cisco Wholesale Voice Solution:

- Chapter 5, "Service A: Minutes Aggregation and Resale (Including ASP Termination)"
- Chapter 6, "Service B: Card Services (Prepaid and Postpaid)"

The following issues are presented as they apply to those services:

- Dial Plan
- Billing and Settlement
- Security
- Prompting

Service A: Minutes Aggregation and Resale (Including ASP Termination)

This chapter provides basic design information related to Service A, Minutes Aggregation and Resale (Including ASP Termination). For the background required to understand the issues discussed, refer to Chapter 2, "Solution Architecture."

For Service A, five templates cover the five call topologies introduced in IP Interconnect Variations, page 4-3:

- Template A1: TDM-to-TDM Call Topology
- Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect
- Template A3: TDM-to-IP-Based Interconnect with OSP
- Template A4: IP-to-IP-Based Interconnect (Transit Network) with DGK
- Template A5: IP-to-IP-Based Interconnect (Transit Network) with OSP

These templates are for simple carrier-interconnect scenarios. In all cases, basic H.323 fault tolerance is achieved as in Fault Tolerance, page 4-8.

For special issues related to limited egress carrier-sensitive routing and interconnect to Clarent-based clearinghouses, see Chapter 7, "Deploying Service Options."

Template A1: TDM-to-TDM Call Topology

Dial Plan

This application uses the basic large-scale H.323 dial plan (gatekeeper core) concept as discussed in Basic Dial Plan, page 4-7.

Billing/Settlement

The wholesale provider dedicates separate GWs for each TDM interconnect partner. The billing system is provisioned to identify carriers by means of originating and terminating GW IP addresses. This allows the wholesaler to generate appropriate CDRs to provide settlement.

Security

Calls in this template type are all intradomain calls. Security may be accomplished as discussed in Security, page 4-9.

Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect

Dial Plan

The basic large-scaleH.323 dial plan concept as discussed in Basic Dial Plan, page 4-7, is still used. To interconnect wholesaler POPs with IP interconnect partners, the peering DGKs simply require additional LRQ route statements to direct certain destination patterns between the wholesaler and the IP-based interconnect partner. Routes between IP interconnect partners and wholesaler POPs are easily added and modified in the DGK, as the rest of the network remains untouched.

Billing/Settlement

In this case, the wholesaler owns only one of the GWs in the conversation, either the originating or the terminating GW, depending on the call direction. The wholesaler's billing application must be able to extract enough information from one side of the call to generate a CDR. This requires correlating either source or destination IP addresses with a particular IP interconnecting carrier, depending on the call direction. To bill the interconnecting customer accurately, the wholesaler's billing system must maintain a database of this information. For calls sourced from ASPs, the list of possible originating IP addresses is typically limited to a few call-signaling proxy servers. However, for ITSPs with many GWs or PC clients, this list can be extensive. The list may be reduced if the ITSP forgoes performance and uses GKRCS. Once carrier identification issues are solved, AAA billing and settlement are done on the GWs.

Alternatively, the originating ITSP or ASP can include a mutually recognized carrier ID (such as prepended ANI) in the H.323 SETUP message. The terminating GW will then include this information in the AAA record. The billing application may be provisioned to recognize this carrier ID and associate it with an originating carrier. However, this implies a trusting relationship between service providers.

Security

Security may be accomplished as discussed in Security, page 4-9. However, this requires the wholesaler to share a database of all gateway user IDs and passwords with all IP-based interconnecting partners. This implies a trusting relationship with partnering carriers.

Template A3: TDM-to-IP-Based Interconnect with OSP

Dial Plan

There are two methods an OSP-based interconnect partner may use to connect into the wholesaler's network, designed as discussed in Basic Dial Plan, page 4-7. This is by implementing OSP either directly on the GW, or through a back-to-back OSP interconnection zone.

From a call routing perspective, OSP is most readily accepted into the network if an OSP interconnection zone is used. This interconnection zone consists of back-to-back GWs. One GW handles the RAS side of the call and the other the OSP side of the call. From the perspective of the DGK, this simply looks like another TDM zone managed by a local GK. The DGK simply adds LRQ routes to the OSP interconnect-zone GK for specific destination patterns serviced by that OSP interconnect partner.

Provisioning requirements for the GWs within this OSP interconnection zone are only slightly different from requirements that are in a normal wholesaler TDM POP. The OSP-side GW is configured to interface with the OSP server. The RAS-side GW is provisioned like a normal POP RAS GW. The back-to-back GWs are then configured to send all calls received over IP out TDM interfaces to the opposite GW, using single-stage dialing. This method of OSP interconnect isolates provisioning tasks to the back-to-back GW pair, the local hopoff GK configuration, and an added LRQ route in the DGK. The rest of the network is unaffected.

If OSP is implemented without an interconnect zone, dial plan provisioning increases dramatically in order to support OSP directly on the GWs. Separate dial peers are needed on all POP GWs to send calls to the OSP server for route resolution, instead of through RAS. The wholesaler may provision dial peers on the GWs to send calls to the OSP server for specific destination patterns.

For example, if the provider knows that all calls to Australia need to be terminated by OSP, the wholesaler may insert a dial peer into its GWs that sends all calls beginning with a "61" country code to an OSP session target. However, any changes to the OSP dial plan require modification to the dial peers on all GWs in the network. This becomes a large burden for a wholesaler with many GWs.

The wholesaler may choose to configure the GW with rotary dial peers, instead of explicit patterns, to handle OSP-based interconnects . Although this may reduce the dial plan's sensitivity to changes, it still requires additional dial-peer provisioning to support failover. In this case, GWs are configured to try to terminate the call within their own administrative domain, first through RAS. If RAS offers no termination possibilities, either by explicit ARJ or RAS timeout, the GWs may fall back to a secondary dial peer to reoriginate the VoIP call through OSP.

For instance, the GW is provisioned with two dial peers having identical destination patterns. This is typically a generic pattern (any nonlocally supported number). One dial peer points to session-target RAS and the other points to session-target settlement. The RAS dial peer is given a higher priority than the settlement dial peer, so that the former is always attempted first. In the event that the RAS dial peer fails, the GW then attempts to send the call to an OSP server through the secondary dial peer.

This method reduces the amount of maintenance of OSP dial peers that is required to accommodate dial plan changes, but adds postdial delay to all OSP-based interconnect calls.

Billing/Settlement

In any OSP implementation, the OSP server collects usage information and generates CDRs. This usage information is extracted directly from the GWs registered to the OSP server, regardless of whether they are functioning as back-to-back GWs or as normal POP GWs.

The wholesaler may also send duplicate records to an AAA server for internal accounting. These AAA CDRs may be used to cross-check any settlement issues with the OSP provider. As an option, the wholesaler may employ a mediation application to automate this process.

Security

If OSP is performed directly on the terminating GW, intradomain security continues to use (optionally) Cisco access lists, as in Template A2. Interdomain security uses OSP H.235 tokens for *registration only*, with the caveats to the dial plan as noted in Security, page 4-9. If a back-to-back GW zone is used, the OSP token management is offloaded from the wholesaler's POP GWs, and is handled instead by the OSP GW in the back-to-back zone. The OSP GW in the back-to-back pair supports the H.235 OSP tokens, whereas the RAS GW (optionally) implements Cisco access lists. This use of the back-to-back OSP transit zone makes it possible to sidestep the security caveats previously mentioned in the direct method.

Template A4: IP-to-IP-Based Interconnect (Transit Network) with DGK

Dial Plan

Interconnections between IP-based service providers are sent to a back-to-back GW transit zone. Each IP interconnecting partner has a dedicated transit zone. If both partners are interconnected through a DGK peering relationship, this adds complexity to the large-scale H.323 dial plan architecture discussed in Basic Dial Plan, page 4-7. The dial plan must be altered to provide dedicated ingress and egress DGKs to route calls properly through the wholesaler network. IP interconnect from one carrier using DGK peering and an OSP-based interconnection partner using a back-to-back OSP interconnection zone is accomplished in essentially the same way as for Template A2.

Billing/Settlement

The back-to-back GW provides a point in the call-signaling path from which the wholesaler may gather accounting information. Billing can be done from the back-to-back GW in the same manner as described as in the simple interconnect method of the TDM-to-TDM scenario in Template A1.

Security

The back-to-back GW zone also allows the wholesaler to obscure an originating ITSP carrier's information from the terminating ITSP carrier, if desired. Calls sent into a terminating ITSP (call it ITSP B) look as if the wholesaler sourced them. ITSP B has no idea that the originating ITSP (call it ITSP A) sourced the call.

Gateway IDs and passwords still must be shared among the wholesaler and interconnecting partners. However, the back-to-back GW allows the wholesaler to isolate interdomain security information between SPs. That is, ITSP A does not need to know ITSP B's security information, and vice versa, for the two to complete calls between each other.

Template A5: IP-to-IP-Based Interconnect (Transit Network) with OSP

Dial Plan

This extends the method described in Template A3 to include sending calls to another OSP provider through another back-to-back GW zone or another DGK-based service provider, depending on LRQ routing entries in the DGK.

Billing/Settlement

Billing is between OSP providers and is done just as discussed in Template A3, but for two OSP back-to-back GW zones. The originating zone provides settlement CDRs for the originating carrier, and the terminating zone provides settlement CDRs for the terminating carrier. If the call is instead sent to a DGK interconnect, then AAA RADIUS records are used on that side. If a mediation application is used, the AAA may be reconciled with the OSP usage records.

Security

Security is accomplished as in Template A3.

CHAPTER 6

Service B: Card Services (Prepaid and Postpaid)

This chapter provides basic design information related to Service B, Card Services (Prepaid and Postpaid). For the background required to understand the issues discussed, refer to Chapter 2, "Solution Architecture."

For Service B, five templates cover the five call topologies introduced in IP Interconnect Variations, page 4-3:

- Template B1: TDM-to-TDM Call Topology
- Template B2: TDM-to-IP Call Topology Using DGK-Based IP Interconnect
- Template B3: TDM-to-IP-Based Interconnect with OSP
- Template B4: IP-to-IP-Based Interconnect (Transit Network) with DGK
- Template B5: IP-to-IP-Based Interconnect (Transit Network) with OSP

In all cases, basic H.323 fault tolerance is achieved as in Fault Tolerance, page 4-8. For special issues related to limited egress carrier-sensitive routing and interconnect to Clarent-based clearinghouses, see Chapter 7, "Deploying Service Options."

Template B1: TDM-to-TDM Call Topology

Dial Plan

Card services typically affect dialing habits by employing two-stage dialing. Aside from this, dial plans remain basic. Once inside the wholesaler network, the call may either be terminated at one of the wholesaler's POPs or sent to another service provider through a TDM hopoff, using the basic large-scale H.323 dial plan architecture as discussed in Basic Dial Plan, page 4-7.

Billing/Settlement

Card services for TDM-based interconnecting partners are supported on the wholesaler's originating GW. AAA-based billing is done on the GWs and is settled as discussed in Billing/Settlement, page 4-6. However, in order to offer prepaid services, the billing server must interact in real time with the AAA server.

Security

An IVR script running on the originating GW performs user authentication. This IVR script interacts with an AAA RADIUS security server. Call-level security may be optionally implemented through Cisco access lists, as discussed in Security, page 4-9.

Prompting

In order to support branding requirements, the wholesaler must be able to identify the necessary IVR script for the carrier. Different call scripts may be invoked, depending on the supplied DNIS. If desired, prompts may be stored remotely on a TFTP server.

Template B2: TDM-to-IP Call Topology Using DGK-Based IP Interconnect

Dial Plan

For card services provided to TDM interconnect partners, this has the same considerations as outlined in Template B1. However, the wholesaler may want to provide card services for IP interconnecting partners as discussed in Service B: Calling Card Services (Prepaid and Postpaid), page 1-5. In this case, the wholesaler may route incoming VoIP calls directly to the terminating GW as normal and then implement the IVR.

Alternatively, the GKs and DGKs may be provisioned to route the call first to a back-to-back GW for IVR services, based on the end user's dialing a specific access number. The DGK knows to send calls destined to this access number to a particular IVR zone consisting of back-to-back GWs. The local GK is provisioned to send calls destined to this access number to a designated ingress-only GW of the back-to-back GW pair. The egress GW is explicitly given a gw-priority of 0 to avoid sending calls through the back-to-back GW in the reverse direction.

The ingress back-to-back GW is provisioned to pass this call through TDM to the egress GW. The egress GW then applies the required IVR script, based upon the DNIS received. The egress GW collects the desired destination pattern and reoriginates the call into the H.323 network, just as with a normal TDM POP.

Billing/Settlement

AAA-based billing is done on the GWs as discussed in Billing/Settlement, page 4-6. However, in order to offer prepaid services, the billing server must interact in real time with the AAA server. For back-to-back GW scenarios, billing is done on one of the GWs as if it were a normal TDM POP.

Security

Security is accomplished in the same manner as described in the simple interconnect application above. Added security is provided by the IVR script in authenticating IP-based users—either before the call enters the wholesaler's network (as with the back-to-back implementation), or at least before the call is completed through the wholesaler network (as with the terminating GW implementation).

Prompting

Prompting for TDM interconnects is the same as in Template B1. In order to support the proper welcome announcements and local languages that are required for branding in IP interconnections, the wholesaler must be able to identify the source carrier before authenticating the user.

Where IVR is implemented directly on the terminating GW, the called number is supplied by the end user and is routed to the destination. It is unreliable to identify the originating carrier on the basis of dialed DNIS. Modifications may be made to ANI, but this is also unreliably enforced on originating PC endpoints. Therefore, multiple branding is not supported in this implementation for IP interconnect partners.

For IP interconnects front-ended with a back-to-back GW, the wholesaler may support branding services to individual carriers by providing separate access numbers, which PC users dial to reach various back-to-back GW zones. For example, carrier A is given a special destination number to dial into its back-to-back GW IVR pool.

Template B3: TDM-to-IP-Based Interconnect with OSP

Dial Plan

Dial plans may be administered in a similar fashion as discussed in the card services application in Template B2. However, in this case, front-ending IVR calls do not require routing to separate back-to-back GW IVR zones. IVR services may be performed directly on the interconnecting OSP back-to-back GW pair.

Billing/Settlement

Billing is done in the same manner as for Template B2.

Security

Security is implemented in the same manner as in Template A2 (Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect, page 5-2). Added security is provided by the IVR script in authenticating IP-based users either before the call enters the wholesaler network (as with the back-to-back GW implementation), or at least before the call is completed through the wholesaler network (as with the terminating GW implementation).

Prompting

Prompting is implemented in the same manner as in Template B2. For OSP interconnects using a back-to-back GW zone, the IVR services may be implemented on the RAS side GW as if it were a normal POP GW.

Template B4: IP-to-IP-Based Interconnect (Transit Network) with DGK

Dial Plan

The wholesaler may want to provide card services for IP interconnecting partners by using a back-to-back GW IVR zone as the front-ending application. See Template A2 (Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect, page 5-2).

Billing/Settlement

Billing is done on one of the GWs as if it were a normal TDM POP. AAA-based billing is done on the GWs as discussed in Billing/Settlement, page 4-6.

Security

Security is accomplished as in Template A4 (Template A4: IP-to-IP-Based Interconnect (Transit Network) with DGK, page 5-4). Added security is provided by the IVR script in authenticating IP-based users before the call traverses the wholesaler’s network in the back-to-back GW.

Prompting

Prompting is done as in Template B2. The back-to-back GW essentially operates as the front-end application.

Template B5: IP-to-IP-Based Interconnect (Transit Network) with OSP

Dial Plan

The wholesaler may want to provide card services for OSP-based IP interconnecting partners by using a back-to-back GW zone, as discussed in Template B2.

Billing/Settlement

Billing is done on one of the GWs as if it were a normal TDM POP, as in Template B2.

Security

Security is accomplished as in Template A5 (Template A5: IP-to-IP-Based Interconnect (Transit Network) with OSP, page 5-5). Added security is provided by the IVR script in authenticating IP-based users before the call traverses the wholesaler's network in the back-to-back GW.

Prompting

Prompting is done as in Template B2 on the back-to-back GW.

Deploying Service Options

This chapter presents various issues—dial plan, billing/settlement, and security—as they relate to feature enhancements to Services A and B. The following service options are discussed:

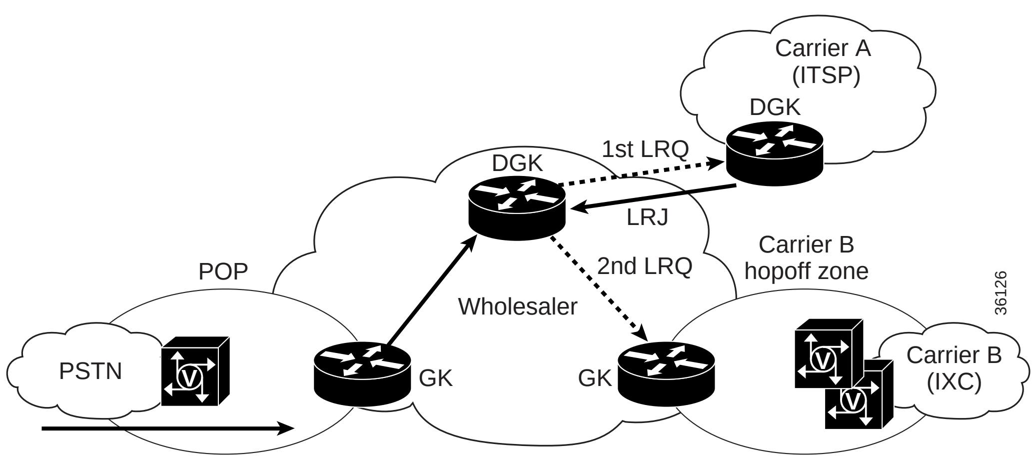
- Limited Egress Carrier-Sensitive Routing
- Interconnect to Clarent-Based Clearinghouses

Limited Egress Carrier-Sensitive Routing

The DGK can make limited egress CSR decisions by using the sequential LRQ feature, which is available to the applications using DGK routing. Generally speaking, this means any TDM partners and DGK peering partners, but also includes any OSP partners where an OSP interconnection zone is used (as opposed to a direct implementation on the wholesaler's GWs).

In this CSR application, the sequential LRQ feature is used to route a call to different carriers, each of whom supports a different destination. For example, the wholesaler may provision its GKs to route certain destination patterns to carrier A first. If carrier A (an Internet telephony service provider, or ITSP) is unavailable, as a result of, say, a location request reject (LRJ) or LRQ timeout, the wholesaler may decide to route the call to carrier B (an interexchange carrier, or IXC), then to carrier C, and so on. Figure 7-1 illustrates this application.

Figure 7-1 Limited Egress CSR Using Sequential LRQs



Considerations

Note the following restrictions on egress CSR:

- The selection of egress carrier is independent of the source carrier.

  The GK routes calls on the basis of DNIS. The wholesaler statically configures a list of possible egress carriers, and these are tried in order. Routing decisions are not based on which carrier sourced the call. For example, the fact that carrier A sourced the call does not determine that carrier A will be selected for the termination.

- Each destination carrier must be contained in its own zone.

  For ITSP carriers, this is fairly simple. Interconnecting ITSPs are seen as single remote zones to which the wholesaler DGK sends LRQ messages. For interconnecting TDM carriers, this implies the following:

  – that the GWs capable of sending calls to the carrier are grouped into their own hopoff zone (which is managed by a GK)

  – that multiple carriers are *never* supported by a single GW

- Dynamic routing decisions are not supported.

  The order of sequential LRQs is configured statically. Consequently, there is no provision for percentage-based routing, maximum minute cutoffs, and the like. Egress carriers are always chosen from a statically configured list of routes. If the DGK determines that an OSP interconnection zone handles a route, it is possible that the OSP server returns a terminating GW on the basis of advanced routing logic (if so provisioned). For example, the OSP server may dynamically select a least-cost, terminating carrier on the basis of time of day or best voice quality.

Applicability: Templates A1 and B1

The following applies to Templates A1 and B1.

Dial Plan

In addition to the basic large-scale H.323 dial plan concept as discussed in Basic Dial Plan, page 4-7, additional LRQ entries may be inserted into the DGK. to try different TDM interconnect zones sequentially in a limited egress CSR application.

Billing/Settlement

Billing and settlement are done as in the simple carrier interconnect application discussed in Billing/Settlement, page 5-2.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

Applicability: Templates A2 and B2

The following applies to Templates A2 and B2.

Dial Plan

This is the same as discussed in the TDM-to-TDM call topology (see Template A1: TDM-to-TDM Call Topology, page 5-1), except that sequential LRQ entries into the DGK may include IP-based carriers with which a DGK peering relationship is made.

Billing/Settlement

Billing and settlement are done as in the simple carrier interconnect application discussed in Billing/Settlement, page 5-2.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

Applicability: Templates A3 and B3

The following applies to Templates A3 and B3.

Dial Plan

The dial plan is administered in much the same way as described in the simple interconnect application (see Dial Plan, page 5-1). However, in this case the wholesaler may attempt different terminating carriers for a call. Actual methods depend on whether OSP is implemented directly on the GW or through an OSP interconnection zone.

For OSP implementations directly on the POP GWs, the OSP server may be provisioned to return the best egress carrier to the originating GW. However, the Cisco Wholesale Voice Solution does not address any specific intelligent provisioning on the OSP server.

If OSP is done by means of a back-to-back GW zone, then the LRQ method discussed in Applicability: Templates A2 and B2, page 7-3, may be used on the DGK to select a terminating carrier sequentially. These terminating carriers may include multiple OSP interconnect zones. If the selected terminating carrier is an OSP back-to-back zone, then the OSP GW can additionally receive a preferred terminating GW from the OSP server on the basis of OSP server provisioning. Again, the Cisco Wholesale Voice Solution does not address any specific, intelligent provisioning on the OSP server.

Billing/Settlement

Basic AAA-based billing is done from the back-to-back GWs as discussed in Billing/Settlement, page 4-6. The AAA record collected from the ingress GW contains information about the originating IP carrier, whereas the AAA record collected from the egress GW contains information about the terminating carrier. The billing application can then generate a CDR for the call and the wholesaler may settle between carriers as usual.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

Applicability: Templates A4 and B4

The following applies to Templates A4 and B4.

Dial Plan

Additional sequential LRQ entries are inserted into the DGK to point to various other IP interconnection partners. These IP interconnect partners may be other DGK-based interconnection partners through a back-to-back GW zone, or an OSP-based interconnect partner through a back-to-back OSP interconnection zone.

Billing/Settlement

Basic AAA-based billing is done from the back-to-back GWs as discussed in Billing/Settlement, page 4-6. The AAA record collected from the ingress GW contains information about the originating IP carrier, whereas the AAA record collected from the egress GW contains information about the terminating carrier. The billing application can then generate a CDR for the call and the wholesaler may settle between carriers as usual.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

Applicability: Templates A5 and B5

The following applies to Templates A5 and B5.

Dial Plan

This extends the method described in Template A3: TDM-to-IP-Based Interconnect with OSP, page 5-3, to include sending calls to another OSP provider through another back-to-back GW zone or another DGK-based service provider, depending on LRQ routing entries in the DGK.

Billing/Settlement

Billing is between OSP providers and is done as in Template A3: TDM-to-IP-Based Interconnect with OSP, page 5-3, but for two OSP back-to-back GW zones. The originating zone provides settlement CDRs for the originating carrier and the terminating zone provides settlement CDRs for the terminating carrier. If the call is instead sent to a DGK interconnect, then AAA RADIUS records are used on that side. A mediation application may be used to reconcile the AAA with the OSP usage records.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

Interconnect to Clarent-Based Clearinghouses

In this application a back-to-back GW must be used, because Clarent does not support either DGK LRQ messaging or OSP. This architecture is very similar to that of using back-to-back GWs to interconnect OSP partners, except that here the relationship is an H.323 GW to a GK, instead of to OSP.

One of the back-to-back GWs registers to a Clarent GK in the Clarent-based service provider's network, and the other registers to a Cisco GK in the wholesaler's network.

Considerations

Note the following limitations where Clarent-based interconnect is used.

- IP-to-IP Interconnect

  The use of back-to-back GWs enables Clarent-based interconnect partners to exchange traffic not only with wholesaler TDM-based interconnects, but also with other IP-based interconnect partners. Those partners may be either DGK- or OSP-based. It may be necessary to modify the dial plan architecture to support DGK-based IP carrier interconnects.

- Interoperability Considerations

  To interconnect with Clarent-based networks, H.323 interoperability must be supported between Cisco GWs and Clarent GKs. Currently, only voice-bearer interoperability is supported for G.711, G.723.1, and G.729 codec types.

- Voice Quality

  Because of tandem compression, back-to-back GWs impair voice quality.

Applicability: Templates A2 and B2

The following applies to Templates A2 and B2.

Dial Plan

The impact of this application on the dial plan closely models that of the simple carrier interconnect application discussed for Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect, page 5-2. The basic large-scale H.323 dial plan concept as discussed in Basic Dial Plan, page 4-7, is still used. However, a back-to-back GW zone is added to interconnect with the Clarent-based clearinghouse. The back-to-back GW is provisioned so that one GW registers to the Clarent GK; the other GW registers to the local GK, which manages the Clarent interconnect zone within the wholesaler network.

The DGK simply requires additional LRQ route statements to direct certain destination patterns to the Clarent-based clearinghouse through a zone consisting of back-to-back GWs. To the local GK that manages the Clarent interconnect zone, this looks just like a normal TDM POP and has all the characteristics of the simple interconnect application discussed in Template A1: TDM-to-TDM Call Topology, page 5-1.

Billing/Settlement

The wholesaler may collect AAA RADIUS records from one of the back-to-back GWs in the same manner as described in the simple carrier interconnect above. These may be used to generate a periodic bill to the Clarent-based interconnect partner. Bills received from the Clarent-based interconnect partner may be verified by comparing these CDRs against those generated by Clarent Command Center before settling the bill.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

Applicability: Templates A4 and B4

The following applies to Templates A4 and B4.

Dial Plan

The dial-plan is similar to that discussed for Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect, page 5-2. However, the Clarent-based clearinghouse interconnect zone may reach other IP-based networks through the DGK. These interconnecting IP-based networks may be direct DGK peering partners or OSP-based interconnects through an OSP interconnection zone.

Billing/Settlement

Billing is done as in Template A2: TDM-to-IP Call Topologies Using DGK-Based IP Interconnect, page 5-2.

Security

Security options are the same as for the simple carrier interconnect application discussed in Security, page 5-2.

APPENDIX A

Open Settlements Protocol (OSP) Clearinghouse Solution

Overview

For packet telephony service providers, terminating long-distance calls has inherent risks. With traditional circuit-switched TDM networks, a relative handful of carriers negotiated bilateral agreements for call interconnections. However, with the emergence of voice-over-IP (VoIP) carriers, the task of ensuring compatibility and interconnections with multiple TDM carriers can be difficult, from both a business and technical standpoint. Another risk is the lack of reliability of many new long-distance service providers. New entrants in this extremely competitive arena must find a way to ensure that call termination agreements are limited to reliable business partners.

Cisco VoIP gateways and clearinghouse solutions offer an alternative. By taking advantage of Open Settlements Protocol (OSP) support in Cisco devices, packet telephony service providers can employ reliable third parties to handle VoIP call termination while leveraging the bandwidth efficiencies and tariff arbitrage advantages that are inherent in IP. Thus the clearinghouses can serve as both a technical and business bridge for VoIP service providers. By signing on with such an organization and using OSP, VoIP carriers can extend service beyond the boundaries of their own network and immediately access the entire clearinghouse network of affiliated service providers.

The benefit is reduced barriers for VoIP competitors in the lucrative long-distance market. By using OSP—the only standard IP interface for VoIP clearinghouse functions—service providers have to do business with only a single settlements provider. As a result, there is no need to negotiate separate agreements with carriers in multiple countries, meet varied technical requirements for interconnection, make repeated arrangements for call accounting, or establish several credit accounts. With the OSP standard, a single clearinghouse can do it all.

At the same time, the OSP clearinghouse solution virtually eliminates the risk in doing business with new service providers with limited credit history—or with carriers in countries subject to currency fluctuations. In addition, it gives virtually every VoIP provider the worldwide calling reach they require.

How OSP Clearinghouse Solution Works

Traditionally, interconnecting carriers calculated settlements based on minutes used in circuits exchanged between their switches, often exchanging Signaling System 7 (SS7) information as well as voice paths. Call authorization was based simply on the physical demarcation point—if a call arrived, it was deemed “authorized.” This required a stable business relationship, except in the case of international traffic, where third-party wholesale carriers often provided such services.

VoIP service providers have had to adapt to such arrangements by terminating calls on the PSTN and reoriginating the call on a circuit switch. However, such an approach limits the cost-effectiveness of New World packet telephony. Even interconnection between VoIP networks was problematic—solutions were usually tightly integrated with individual vendors' proprietary and nonstandard implementations of H.323 protocols.

OSP avoids this problem by using a standard protocol approved by the European Telecommunications Standards Institute Internet Protocol Harmonization over Networks (ETSI TIPHON) organization. By allowing gateways to transfer accounting and routing information securely, this protocol provides a common ground for VoIP service providers. That way, third-party clearinghouses with an OSP server can offer call authorization, call accounting, and settlement—including all the complex rating and routing tables necessary for efficient and cost-effective interconnections. It works as follows.

A user places a call over the PSTN network to a VoIP gateway, which authenticates the user by communicating with a RADIUS server. Next, the originating VoIP gateway attempts to locate a termination point within its own network by communicating with a gatekeeper through H.323 RAS. If there is no appropriate route, the gatekeeper informs the gateway to search for a termination point elsewhere. The gateway then contacts an OSP server at the third-party clearinghouse.

At that point, the gateway establishes an SSL connection to the OSP server and sends an authorization request to the clearinghouse. OSP servers from Cisco providers supply the least-cost and best-route selection algorithms according to the carrier's requirements for cost, quality, and other parameters, selecting up to three routes. The clearinghouse creates an authorization token, signs it with a clearinghouse certificate and private key, and then replies to the originating gateway with a token and three selected routes. The originating gateway uses the IP addresses supplied by the clearinghouse to set up the call.

The terminating gateway accepts the call after validating the token. At the end of the call, both the originating and terminating gateways send usage indicator reports to the OSP server. The usage indicator report contains the call-detail information for the OSP server to provide settlement service between the originating and terminating service providers.

Cisco providers are central to the development of the OSP clearinghouse solution. By combining industry-leading expertise in call accounting and routing with industry-leading Cisco VoIP gateway technology, the solution enables clearinghouses to offer worldwide scope to Internet telephony service providers. All this is made possible by the Cisco OPT (Open Packet Telephony) framework.

OPT: The Cornerstone of OSP Clearinghouse Solutions

Where time-division multiplexing (TDM) networks literally buried services in the circuit switch, Cisco OPT brings them into the open by using a layered framework to separate infrastructure, call control, and services.

The connection layer is an ATM- or IP-based function that establishes and manages bearer connections in response to control messages from the call-control plane. Tightly integrated voice gateways encode and decode voice signals, and the low-latency Cisco packet network supports the QoS that voice services require.

The call-control layer processes call requests and instructs the connection layer to establish the appropriate bearer connection. H.323 call-control standards provide this function today. In the future, OPT will support SGCP/MGCP. All this provides complete interoperability between legacy TDM environments and packet networks.

The service application layer applies the service logic. Using standards-based protocols, third-party application providers are given access to all the capabilities of Cisco packet telephony gateways.

The Role of Alliances

By establishing an open-protocol, standards-based approach to packet telephony, Cisco OPT creates an ecosystem of providers who can develop new applications rapidly—and independently of any particular switch vendor.

OSP Solution Benefits

OSP solutions are based on Cisco VoIP gateways with embedded OSP client software, clearinghouse OSP servers, and optionally, public key certificate authorities.

VoIP service providers who use OSP gain several benefits:

- End-to-end VoIP support
- Cost-effective worldwide calling coverage
- Guaranteed settlement of authorized calls
- Incremental revenue increase by terminating calls from other service providers
- Simplified business and credit relationships
- Outsourced complex rating and routing tables
- Flexibility in selecting appropriate termination points
- Secure transmission using widely accepted encryption for sensitive data

GLOSSARY

For terms or acronyms not listed below, see Internetworking Terms and Acronyms at the following URL:

http://www.cisco.com/univercd/cc/td/doc/cisintwk/ita/index.htm

A

AAA authentication, authorization, and accounting

ALTDGK alternate directory gatekeeper

ANI automatic number identification

ARJ authorization reject

ARP authorization permit

ASP application service provider

ARQ authorization request

B

BHCA busy hour call attempts

C

CAS channel-associated signaling

CDR call detail record

CIC Cisco Info Center

CLI command line interface

CO central office

CSR customer service record

CVM CiscoWorks2000 Voice Manager

D

DGK	directory gatekeeper
DHCP	Dynamic Host Configuration Protocol
DNIS	dialed number identification service
DRQ	disconnect request

E

EMS	element management system
EO	end office
ETSI TIPHON	European Telecommunications Standards Institute Internet Protocol Harmonization over Networks

G

GK	gatekeeper
GKRCS	gatekeeper-routed call signaling
GKTMP	GateKeeper Transaction Message Protocol—a Cisco-proprietary protocol that allows third-party applications to influence the operation of the IOS GK
GW	gateway

H

HSRP	Hot Standby Router Protocol—used to ensure GK fault tolerance

I

IMT	intermachine trunk
IPM	Cisco’s Internetwork Performance Monitor
ISP	Internet service provider
ITSP	Internet telephony service provider
IVR	interactive voice response
IXC	interexchange carrier—a regulated U.S. Class 4 carrier that is often a wholesale customer

L

LEC	local exchange carrier
LCR	least-cost routing

M

MGCP	Media Gateway Control Protocol
MIB	management information base

N

NI-2	National ISDN version 2—a BRI circuit
NMS	network management system
NOA	nature of address

O

OLO	other licensed operator
OSP	Open Settlements Protocol
OSS	operations support system
OPT	Open Packet Telephony

P

PDF	portable document format
PIN	personal identification number
POP	point of presence
PSTN	Public switched telephone network
PTT	Post, Telephone, Telegraph—a government-mandated or -operated national telephony carrier

Q

QoS quality of service

R

R2 A type of CAS used widely in places other than North America

RADIUS Remote-Authentication Dial-In User Service

RAI resource availability indicator

RAS H.225 Registration, Admission, and Status Protocol—spoken between H.323 gateways and their gatekeepers

RLM redundant link manager

RRQ registration request

RTP Realtime Transport Protocol

RTSP Real-Time Streaming Protocol— for controlling the streaming of RTP packets from a storage source

S

SC signaling controller—a Cisco SC2200 signaling gateway that converts SS7 to a backhauled NI-2 protocol to gateways

SGCP Simple Gateway Control Protocol

SLT signaling link termination—a Cisco 2611 machine capable of terminating SS7 at the MTP2 layer and backhauling MTP3 (and up) to the SC2200 or virtual switch controller (VSC)

SNMP Simple Network Management Protocol

SS7 Signaling System 7

SSL secure socket layer

T

TAC Technical Assistance Center

TCL Tool Command Language

TDM time division multiplex

TFTP Trivial File-Transfer Protocol

U

URL uniform resource locator

UDP User Datagram Protocol

URQ unregistration Request

V

VSA vendor-specific attribute—a nonstandard attribute tag used by RADIUS. Cisco has defined many useful VSAs to enhance the gateway CDR format.

VSC virtual switch controller—one of various Cisco machines capable of providing SS7 signaling conversion, and able to control gateways by means of MGCP

W

WIC WAN interface card


INDEX

C

D

E

F

G

H

I

L

M

N

O

P

Q

R

S

X

Z