IP COMMUNICATIONS

VoIP Performance Management and Optimization

A KPI-based approach to managing and optimizing VoIP networks

Adeel Ahmed, CCIE® No. 4574
Habib Madani
Talal Siddiqui, CCIE No. 4280

ciscopress.com

VoIP Performance Management and Optimization

Adeel Ahmed, Habib Madani, Talal Siddiqui

Published by:
Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

ISBN-13: 978-1-58705-528-7

ISBN-10: 1-58705-528-7

Printed in the United States of America

First Printing July 2010

Library of Congress Cataloging-in-Publication Data:

Ahmed, Adeel.
VoIP performance management and optimization / Adeel Ahmed, Habib Madani, Talal Siddiqui.
p. cm.
ISBN 978-1-58705-528-7 (hardcover)
1. Internet telephony—Management. 2. Computer network protocols. I. Madani, Habib, 1969- II. Siddiqui, Talal, 1973- III. Title.

TK5105.8865.A38 2010
621.385—dc22 2010023573

Warning and Disclaimer

This book is designed to provide information about managing and optimizing VoIP networks using a metrics-based approach that relies on collecting, analyzing, and correlating VoIP performance data from various network elements. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an "as is" basis. The authors, Cisco Press, and Cisco Systems, Inc., shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

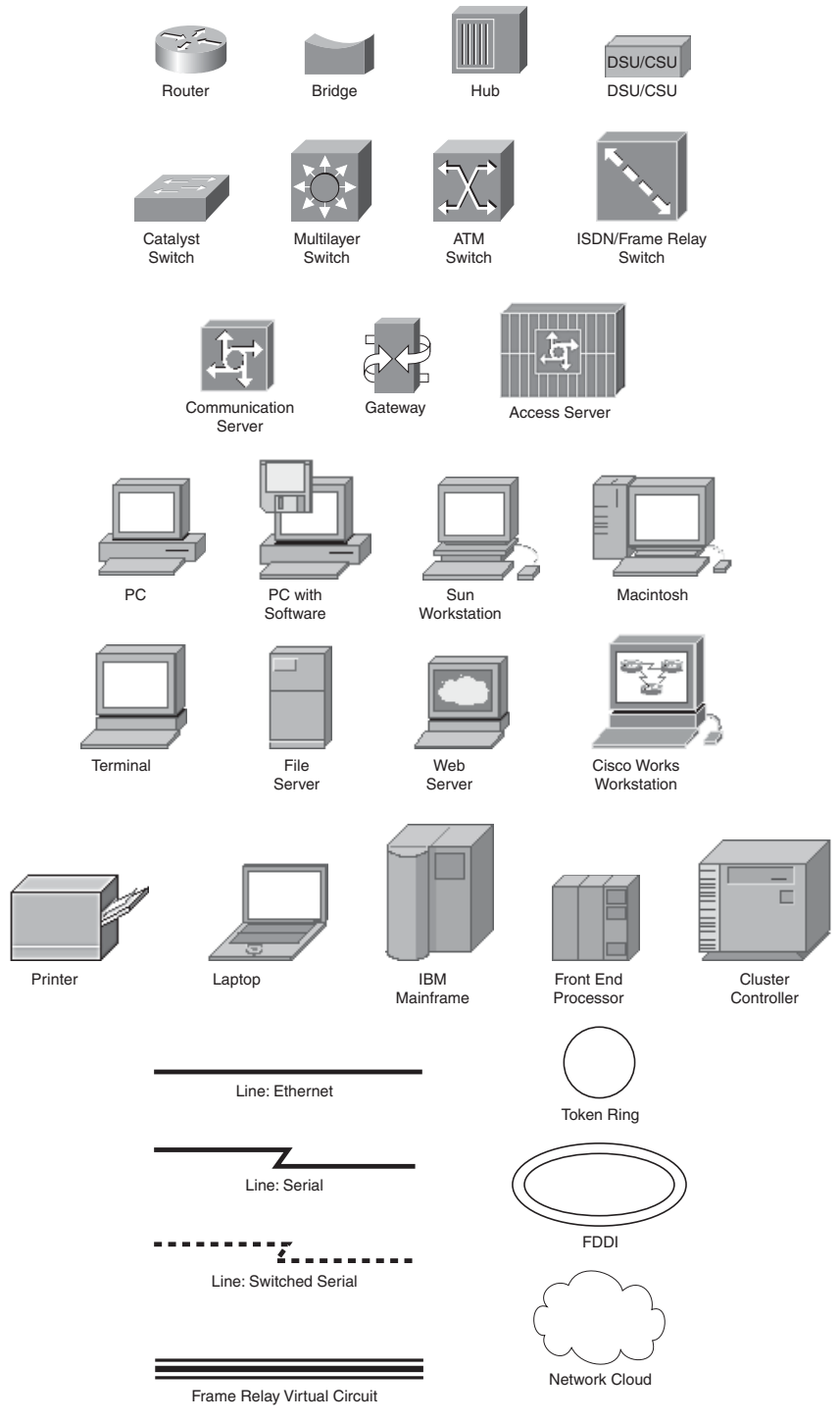
Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Contents at a Glance

Contents

Icons Used in This Book



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a **show** command).
- *Italic* indicates arguments for which you supply actual values.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets ([ ]) indicate an optional element.
- Braces ({ }) indicate a required choice.
- Braces within brackets () indicate a required choice within an optional element.

Foreword

According to a Gartner market share analysis done for Enterprise Unified Communications on June 23, 2009, the total vendor revenue from the entire enterprise unified communications segment in 2008 was $5.1 billion. FierceVoIP quoted ISP-Planet study in its online newsletter on July 28, 2008, saying that the total subscribers in just the United States for the top 10 VoIP service providers are around 18 million and growing in double digits. Comcast came in on top at 5.2 subscribers followed by Time Warner (3.2 million) and Vonage (2.6 million), based on their first-quarter reporting in 2008. This puts great emphasis on managing VoIP for both enterprises and service providers.

Management of a VoIP network is a cyclic process that starts even before VoIP is deployed. The first stage is planning, which includes forming a team, defining the scope of deployment, requirements validation, and assessment of the IP network to determine whether the infrastructure is adequate to support media traffic. The design phase includes comprehensive design-based traffic engineering and validated requirements. It not only covers call-processing servers, remote gateways, and features implementation but also covers changes to the IP network in the form of quality of service design and provisions for network resiliency. It is followed by the implementation phase, which is governed by project management principles and ensures that best practices for deployment are followed for on-time completion.

Implementation also includes a test plan execution and transfer of information to network operations prior to commissioning. The correct execution of these phases ensures minimum problems and decreases the total cost of deployment. Implementation is followed by the operations phase, with continuous optimization to close the loop. This book briefly mentions planning, design, and implementation stages and emphasizes the operations and optimization phases.

First, the hand-off to operations needs to be complete, including remediation of any issues discovered when the postdeployment test plan was executed. All the deployed devices must be discovered by the network management systems. But most important, VoIP can no longer be managed in a silo that is separate from the data network management subteam. This book emphasizes correlating network problems with VoIP-related key performance indicators for faster problem resolution by isolating it and fixing the root cause.

Operational data provides critical feedback for continuous optimization of the network, including its performance and capacity. Optimization is not limited to fine-tuning the traffic engineering process for future growth but also for extending VoIP for the next evolution to collaboration-enabled business transformation.

What is presented in this book is the authors' collective experience and knowledge, working with several other colleagues from Advanced Services, Cisco Remote Operations Service, the product development teams, and most important, Cisco customers, whose feedback was critical in developing best practices for VoIP management and optimization.

Regards,

Talal Siddiqui, Senior Manager, Unified Communication/Collaboration Practice Cisco Advanced Services

Introduction

With the exponential growth of the Internet and an increasing number of VoIP deployments, customers are looking for new ways to manage and scale their networks to meet the growing needs of end users. Customers not only need to fix problems in a timely manner with minimal downtime, but they also need to proactively monitor their networks to fix potential problems before they become service and revenue impacting.

The complexity of an IP network increases with the addition of new services, and as these networks start to scale, managing them becomes a challenge. Customers are looking for new ways to manage their networks and effectively scale these services.

Customers are looking for new techniques and efficient ways to monitor multivendor products in the network and use tools/applications that can scale with the growth of their networks. We got feedback from our customers and VoIP SPs through forums such as SANOG, NANOG, APRICOT, and Cisco Live (formerly known as Networkers) about what they would like to see in a VoIP management book. This feedback can be boiled down to "We want a practical guide with specific details and examples that we can use right away...something that is a desk reference for NOC (Network Operations Center) staff and the network architects."

This book addresses some of the challenges associated with deploying and managing VoIP networks and also provides guidelines on how to optimize these networks.

Goals and Methods

The most important goal of this book is to help define a methodology and framework of collecting, analyzing, and correlating VoIP performance data from various network elements. When correlated in a meaningful way, this data can help network operators identify problematic trends in their VoIP networks, and isolate and fix problems before they become service impacting.

One key methodology in this book is to use a layered approach when troubleshooting VoIP network problems. This helps narrow the scope of the problem in an efficient manner and also helps find the root cause. By quickly identifying the root cause of the problem, the network operator can resolve issues in a timely manner and minimize customer impact.

This book also provides guidelines for optimizing VoIP networks by defining the following:

- What VoIP performance data should be collected from various network elements?
- How to collect VOIP performance data?
- How to use dashboards to analyze and correlate VoIP metrics?
- How to use the VoIP dashboard for trend analysis and capacity planning?

Who Should Read This Book

This book is meant to be used as a guide by network engineers, architects, and operations personnel in managing and optimizing their VoIP networks.

This book also helps network operators troubleshoot VoIP-related issues efficiently and identify root causes to fix problems in a timely manner. However, it does not focus on traces, logs, and debug messages but rather on analyzing trends and correlating network issues to address core issues. This book compliments other Cisco Press publications:

- Kaza, Ramesh and Asadullah, Salman. *Cisco IP Telephony: Planning, Design, Implementation, Operation, and Optimization*. Indianapolis, IN: Cisco Press, February 23, 2005.
- Halmmark, Addis, Giralt, Paul and Smith, Anne. *Troubleshooting Cisco IP Telephony*. Indianapolis, IN: Cisco Press, December 11, 2002.
- Clemm, Alexander. *Network Management Fundamentals*. Indianapolis, IN: Cisco Press, Nov 21, 2006.

How This Book Is Organized

This book discusses some of the challenges faced by service providers and enterprise customers in deploying, managing, and optimizing VoIP in their networks. It provides guidance on how to address voice quality issues and proactively monitor key performance indicators (KPI) to help gauge the health of the VoIP network.

The first part of the book provides an overview of VoIP and key network management concepts. It also discusses a metrics-based approach of managing and optimizing VoIP networks.

The second part of the book concentrates on different VoIP deployment models in SP and enterprise networks, and reviews the common VoIP-related problems in each deployment approach.

Note The first and second parts of the book set the stage for how VoIP is deployed in enterprise and SP networks and discusses the challenges associated with such deployments. You might feel that both these parts of the book are brief and high-level; they do not cover in-depth technology and protocol details. For example, what is DOCSIS and how does it work? How does the Session Initiation Protocol (SIP) work, and what are the various SIP messages? This is by design; it is assumed that you already understand these basics as this information has already been covered in various other texts. The main focus of this book is on managing and optimizing VoIP networks; these concepts are covered in detail in the third part of the book. That is why chapters in the third part of the book are longer and more detailed than the chapters in the first and second parts of the book.

The third part of the book focuses on a proactive approach to diagnosing problems in VoIP networks and fixing these problems before they become service impacting. This part of the book also talks about what tools can be used by customers in gauging the health of their VoIP network and improve network performance. Using performance counters, Call Detail Records (CDR), and Call Agent trace logs, customers can utilize real-time data to gauge the health of their voice network and make capacity-planning decisions before network resources get congested.

Chapters 1 through 8 cover the following topics:

- Chapter 1, "Voice over IP (VoIP) and Network Management": This chapter talks about VoIP media transport and signaling protocols, some common voice quality issues, and their causes. The second half of the chapter discusses network management methodologies such as Telecommunications Management Network (TMN); Fault, Configuration, Accounting, Performance, and Security (FCAPS); and Information Technology Infrastructure Library (ITIL). It also talks about the strategic importance of managing VoIP networks.
- Chapter 2, "A Metrics-Based Approach for Managing the VoIP Network": This chapter highlights the key performance indicators that can be utilized to effectively manage a VoIP network. It encourages the use of a layered approach for isolating localized and systemic issues. It explains how performance data from various network segments and service flows can be used to manage SLAs in a VoIP network.
- Chapter 3, "VoIP Deployment Models in Service Provider Networks": This chapter discusses various VoIP solutions in an SP environment. The deployment models cover scenarios in which broadband SPs provide VoIP service to residential and business customers. These providers own the last-mile connection to end users; they use their infrastructure to not only provide Internet connectivity but also to offer VoIP services using the same infrastructure. Because they own the last-mile connection and the VoIP infrastructure, they can provide better QoS to VoIP traffic and offer high-quality VoIP services.
- Chapter 4, "Internet Telephony": This chapter describes how VoIP is deployed over a publicly shared infrastructure such as the Internet. In such deployment models, the company providing VoIP services might not own the entire network infrastructure, such as the last-mile connection to the end users, which is used for deploying this service. They might use infrastructure, owned by other entities, to provide VoIP as an overlay service by deploying some of their own network components that are required for offering the VoIP service. This deployment model is different from the models discussed in Chapter 3. The VoIP SP is faced with several challenges with providing QoS to VoIP traffic; these issues are also discussed in this chapter.
- Chapter 5, "VoIP Deployment Models in Enterprise Networks": This chapter explains various deployment models that are commonly used in typical enterprise networks, including the fundamental models: central call processing and distributed call processing. It also discusses large-campus deployment schemes.

This chapter discusses the differences in hosted and managed services around Unified Communications solutions. It also presents a brief overview of IP Contact Centers, which are essentially an extended functionality of a Unified Communications solution.

- Chapter 6, "Managing VoIP Networks": This chapter discusses the best practices for planning media deployment over IP networks starting from how to assess the readiness of the network, traffic engineering, high availability, and managing the IP network and its integrated components that process voice and other media transmissions. This chapter also covers the monitoring mechanism available to network administrators and their scope and effectiveness in managing VoIP networks.
- Chapter 7, "Performance Analysis and Fault Isolation": This chapter discusses an approach for proactive monitoring of the VoIP network for performance analysis and fault isolation of problems caused by anomalies in the network. It starts with explaining the VoIP network monitoring aspects including collection, categorization, and correlation of performance counters for both enterprise and service provider networks. It also discusses different ways of gauging the performance of a large-scale VoIP network by looking at various key performance indicators (KPIs).
- Chapter 8, "Trend Analysis and Optimization": This chapter explains the use of VoIP dashboards to monitor and trend performance data from different components in the VoIP network. This trend analysis can help network operators not only establish a baseline but also help with resource optimization and capacity planning by looking at problematic trends in the network, such as resource overutilization and changes in traffic patterns.

Chapter 3

VoIP Deployment Models in Service Provider Networks

This chapter gives you an understanding of how Voice over IP (VoIP) is deployed in service provider (SP) networks. This chapter focuses on describing a use case in which the VoIP infrastructure and the transport and the access are managed by an SP. Chapter 4, "Internet Telephony," focuses on VoIP networks in which only the VoIP infrastructure is managed. Different network components and their functions are described to illustrate how various call functions are implemented to provide voice services to residential and business customers. Figure 3-1 depicts a block architecture of the SP scenarios discussed in this chapter. Here, the service provider also owns the last-mile network access. Later chapters cover scenarios where the SP does not own the access network.

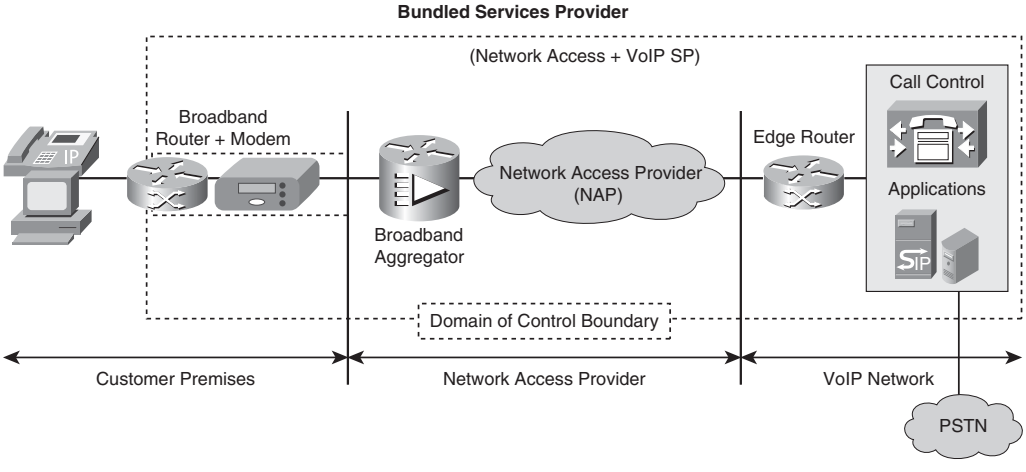


Figure 3-1 *Service Provider Architecture Overview*

This chapter provides a high-level view of the connectivity between different components in a VoIP SP. You learn about the common VoIP networks and the corresponding components. The intention of this chapter is not to provide design guidelines or technology-specific reference material, which is outside the scope of this book, but to offer a collection of metrics from across the various VoIP architectures. As a general note, the acronym *KPI* (key performance indicators) is used throughout the book to refer to key protocol counters or metrics.

This chapter covers various VoIP applications in the SP market; residential application is geared toward providing primary- or secondary-line voice services to SP's residential customers. These customers include existing high-speed data subscribers and new subscribers who are looking at either replacing their current circuit-switched telephone line or adding additional phone lines to their household. This gives SPs a chance to provide bundled services to their customers.

Another application covered in this chapter is Small/Medium Business (SMB) application, which is geared toward business customers. SPs can provide high-speed data and digital voice services to their business customers using their IP infrastructure. For the SMB, using IP infrastructure can be a more cost-effective way of getting voice services as compared to a traditional circuit-switched phone line from the telephone company.

Other applications discussed in this chapter include IP trunks, which are used for traffic offload and public switched telephone network (PSTN) bypass, and Sessions Border Controllers (SBC), which are used for offloading VoIP traffic to the PSTN, network hiding, and voice transcoding.

The latter part of the chapter highlights some of the security-related issues in SP voice networks. These issues include denial of service (DoS) attacks, theft of service, and other issues that are common in existing IP networks today.

The last part of the chapter discusses common issues and problems related to voice in SP networks. Because VoIP is primarily deployed on a converged IP network, it faces many of the same challenges as other data applications, such as failures in the network, routing protocol convergence issues, oversubscription of network resources, and so on. However, because VoIP is more sensitive to things like delay and jitter, it's important to proactively monitor the health of the SP network and prevent network outages or performance degradation that can cause loss of service to its customers. These issues are discussed in more detail in Chapter 6, "Managing VoIP Networks," Chapter 7, "Performance Analysis and Fault Isolation," and Chapter 8, "Trend Analysis and Optimization."

Service Provider Voice Implementation Models

This section goes into the details of different SP voice deployment models. Various network components and their functions are discussed with illustrations. There are two different VoIP implementation models in SP networks:

- **Centralized Switching Model:** In this model, the call-processing functions are controlled by a central entity such as a Softswitch (Call Agent or Call Management Switch [CMS]), which passes call control information to different network elements, sets up and tears down calls, and keeps data records for the calls as Call Detail Records (CDR). The endpoints do not need to have intelligence in regard to initiating or terminating calls; they receive the information from the Softswitch and carry out the necessary call functions.
- **Distributed Switching Model:** In this model, the call-processing functions are distributed to different network elements. A single entity does not control the various call functions. In this model, the endpoints have call intelligence and can initiate and tear down calls without a centralized entity controlling them. The current VoIP SPs are hesitant to go this route, because it makes the end VoIP clients fatter or richer in features and they do not need to subscribe to the SP's premium services. IP Multimedia Subsystem (IMS) is the route that SPs are looking into where presence servers are used to track the end clients.

This chapter primarily focuses on the centralized switching model because most of the current SP deployments are based on this model. The other common distributed switching model is introduced briefly, but it is discussed in more detail in Chapter 4, which also covers some of the current Peer-to-Peer Distributed switching models. The next section covers how the centralized and distributed switching models are deployed in different SP networks.

Residential Applications: Voice over Broadband

In a voice over broadband deployment model, the SP uses the IP infrastructure to provide residential IP telephony services to its customers. An example of such an implementation model is the PacketCable architecture defined by Cable Television Laboratories (CableLabs) PacketCable specifications. The PacketCable specifications define a framework of how VoIP can be implemented over the Data Over Cable Service Interface Specification (DOCSIS)/IP infrastructure. Figure 3-2 provides a high-level overview of the PacketCable architecture. The system uses IP technology and QoS to provide high-quality transport for the VoIP network.

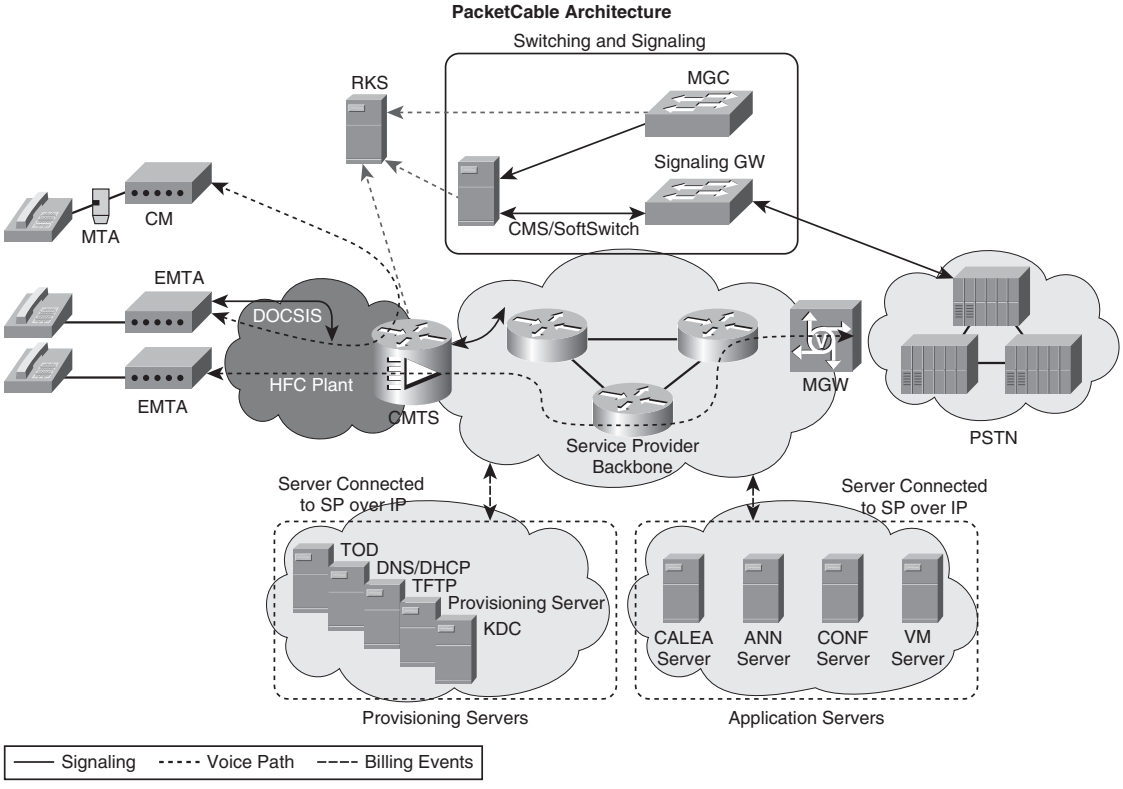


Figure 3-2 *PacketCable Architecture Overview*

The following are some of the key elements of the PacketCable network:

- **Call Management Server (CMS):** The CMS is responsible for providing call control and signaling for the endpoints using Media Gateway Control Protocol/Network-Based Call Signaling (MGCP/NCS) protocol. In a centralized switched model, all the intelligence resides on the CMS, which is responsible for instructing other network elements on their functions.

  The CMS is composed of several logical components, such as Gate Controller (GC), Media Gateway Controller (MGC), Signaling Gateway (SG), and Announcement Controller (ANC). The GC is responsible for quality of service (QoS) authorization and control. The MGC provides call control and signaling for PSTN Media Gateways. The SG communicates call signaling to the PSTN using protocols such as Signaling System 7 (SS7). The ANC interfaces with the Announcement Player (ANP) to play network announcements.

- **Cable Modem Termination System (CMTS):** The CMTS sits at the edge of the network and connects the endpoints to the SP infrastructure such as provisioning servers, CMS, Media Gateway (MGW), and so on over the DOCSIS Hybrid Fiber

Coax (HFC) network. It also allocates resources for voice calls when instructed by the CMS and upon receiving requests from the endpoint.

- **Media Terminal Adapter (MTA)/Embedded-MTA (EMTA):** MTA connects the subscriber equipment, such as a host PC or analog phone, to the SP network over the DOCSIS (HFC) network. It establishes a physical connection with the CMTS and forwards traffic between the SP network and the subscriber equipment. It contains a network interface, radio frequency (RF) interface, CoderDecoder (CODEC), and all signaling and encapsulation functions required for VoIP transport, class features signaling, and QoS signaling.
- **Media Gateways (MGW):** The MGW provides bearer connectivity to the PSTN and is used for off-net calls (when an SP customer calls someone connected to the PSTN, basically an IP-to-PSTN network call).
- **Provisioning Servers:** Figure 3-2 includes a setup of servers; they perform provisioning and billing functionalities. These servers include the Dynamic Host Configuration Protocol (DHCP) server for assigning IP addresses and other network parameters to the endpoints, Domain Name Servers (DNS) for name resolution, Trivial File Transfer Protocol (TFTP) for downloading configuration files to MTAs, and optionally other servers such as syslog server and Ticket Granting Server (TGS), which are used in the PacketCable network.
- **Application Servers:** These servers include voicemail (VM) servers for providing voice mailbox service to subscribers, conferencing servers for audioconferencing service, announcement servers for playing network announcement messages, and Communications Assistance for Law Enforcement Act (CALEA) servers for subscriber wiretapping for law enforcement agencies.
- **Record Keeping Server (RKS):** These are used for billing purposes. They store call detail record information through PacketCable Event Messaging.

Residential gateways in the form of MTA embedded in a cable modem are also known as Embedded Multimedia Terminal Adapters (EMTA). VoIP access is provided at the customer premises. By plugging a standard analog telephone into the MTA device, a user can make phone calls to another Multiple System Operator (MSO) customer directly across the IP network or to anyone outside the SP or MSO network through an MGW.

CMSs and MGCs provide centralized call-control processing by passing control information and setting up connections between residential MTAs. After these connections are established, voice passes directly between gateway endpoints in the form of RTP packet streams, as shown in Figure 3-3. Most connections with the PSTN are through voice bearer trunks with a Media Gateway providing the bearer connections and a Signaling Gateway (SG) providing the signaling connection into the SS7 network. Multi-Frequency/Channel Associated Signaling (MF/CAS) trunks are provided for some specialized requirements, such as Operator Services.

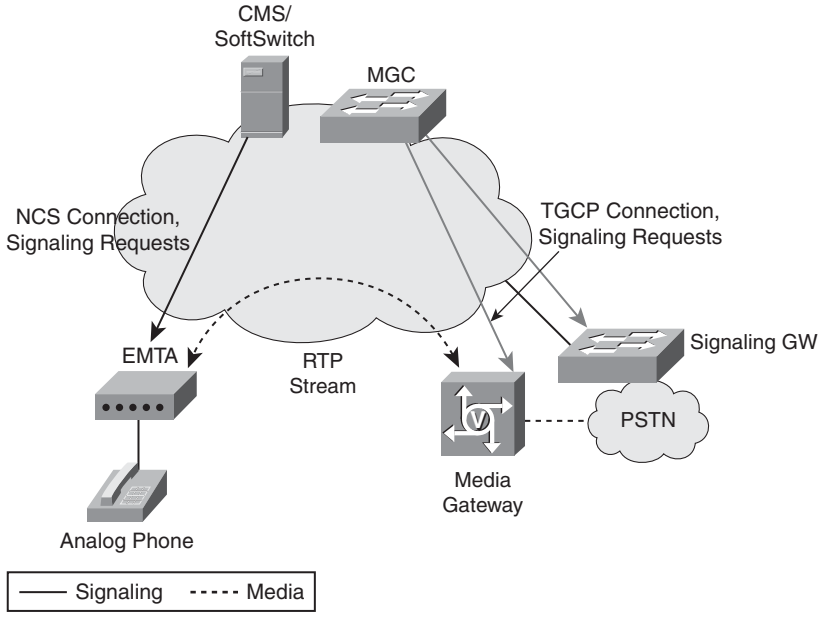


Figure 3-3 *PacketCable Signaling Architecture*

The PacketCable Network-Based Call Signaling (NCS) protocol is used to communicate with the MTA endpoints. The PacketCable Trunking Gateway Call Signaling Protocol (TGCP) is used to communicate with Media Gateways. NCS and TGCP are profiles of the Multimedia Gateway Control Protocol (MGCP), which belongs to the xGCP suite of protocols. These protocols allow a central call control mechanism to control customer premises equipment (CPE) devices for voice services.

Small/Medium Business Applications (Voice over T1/E1/PRI/CAS)

The SP uses small/medium business applications to provide IP-based telephony to SMB customers, typically when only a low number of devices are needed at the customer premises (usually this number is less than 50).

SPs often use an Integrated Access Device (IAD) to provide voice and data service to the customer premises. The IADs are connected to a Local Exchange Carrier (LEC) leased Primary Rate Interface (PRI)/Channel Associated Signaling (CAS) line, which is aggregated through a bigger transport connection like DS3 to the VoIP SP. The CMS residing at the SP is used to provision the IADs.

MGCP is the signaling protocol used to communicate to the IADs, and both signaling and the data ride the same leased line. Figure 3-4 reflects such a topology. The SP also routes PSTN/SS7-bound calls originating at the customer premises through the CMS

that's managing the IAD/MGCP link. The trunking gateway and the SS7 gateway (which can be a Cisco Internet Transfer Point [ITP]) handle the PSTN-bound bearer and signaling traffic, respectively. The CMS acts as a central switching point and is thus an ideal place for collecting key performance indicators (KPI) because it acts as central switching component. The MGCP-based communication counters for announcement servers, trunking gateway and IADs, along with SIP-based communication counters for voicemail server and PSTN-related SS7 signaling (SIGTRAN) protocol counters make up the KPIs for the Small Business model.

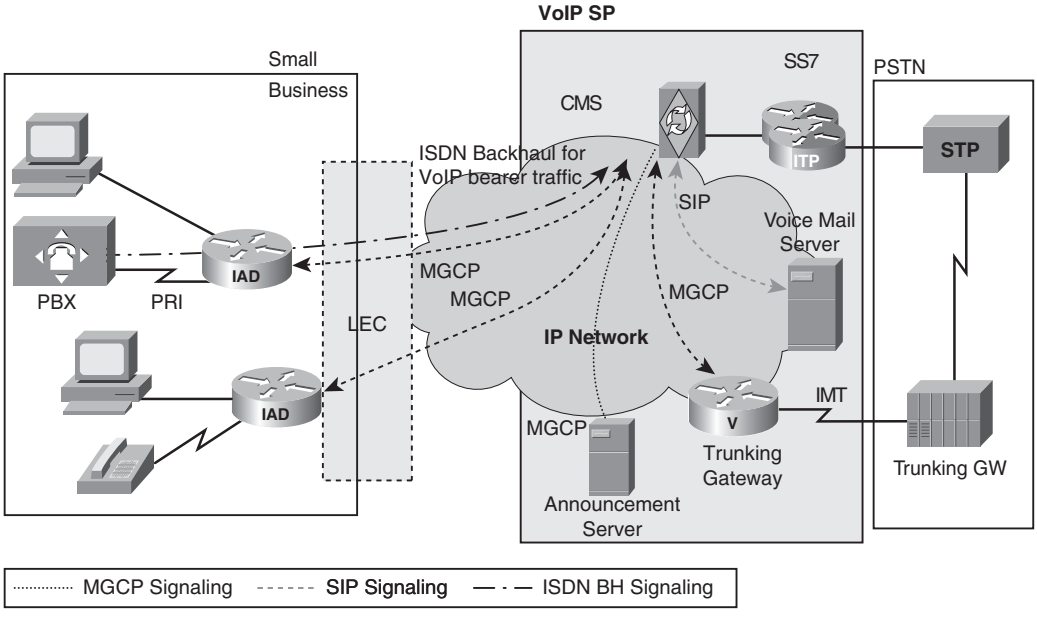


Figure 3-4 *Small/Medium Business Deployment Architecture*

Integrated Access Device (IAD) is also one of the key elements for the small-business network. It can be a collection point for MGCP-related metrics. An IAD is a device used to multiplex and demultiplex traffic in the customer's premises. The IAD is used primarily to route traffic and signaling over to a single T1 line or to an ISDN PRI trunk. This is also called a voice gateway that utilizes E1 lines in the case of non-U.S. markets.

IP Trunks

The SP market is converging to IP, but the PSTN is still the prevalent infrastructure and will be utilized for a while. However, more SPs use the IP network where possible. They achieve this by placing a trunking VoIP switch, which provides Class IV or Tandem switch functionality, to most commonly address two needs. First, it serves as a long-haul SIP trunking switch to carry traffic between SPs of different regions. Second, it acts as a PSTN bypass and an inter-SP trunk interconnect to offload long-distance traffic. The architecture presented in Figure 3-5 touches on various technologies, but the discussion in this section is focused on signaling protocol–related metrics.

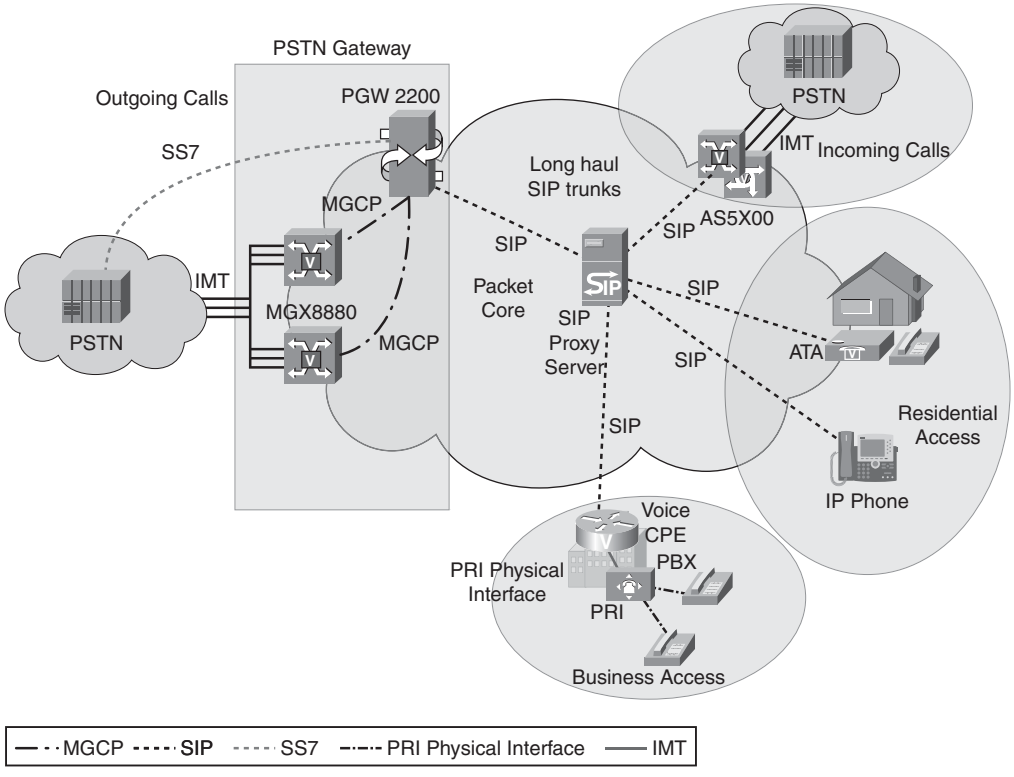


Figure 3-5 *IP Trunk Deployment Architecture*

Some of the key components of the IP trunk architecture are as follows:

- **The Cisco PSTN Gateway (PGW):** The Cisco PGW 2200 is a carrier-class call agent that performs the signaling and call-control tasks (such as digit analysis, routing, circuit selection, and more) between the PSTN and the IP infrastructure. PGW is also called a trunking switch, and it performs Class IV–type functionality.
- **Analog Telephone Adapter (ATA):** The Cisco ATA 186 is a handset-to-Ethernet adapter that turns traditional telephone devices into IP devices, which enables the analog phones to be connected to an IP network. Customers can take advantage of the many new IP telephony applications by connecting their analog devices to Cisco ATAs.
- **SIP Proxy Server:** The Cisco SIP Proxy Server is a call control engine that enables SPs to build scalable, reliable VoIP networks today. Based on the SIP, the Cisco SIP Proxy provides a full array of call-routing capabilities to maximize network performance in both small- and large-packet voice networks.
- **Cisco Access Servers:** The Cisco access gateway provides universal port data, voice, and fax services on any port at any time. It is used as a common gateway for terminating IP trunks that carry VoIP and other types of traffic. It can be a collection point

for signaling and media metrics. Cisco MGX 8850 and AS5400 are examples of the access gateways and are depicted in Figure 3-5.

IP phones also communicate through SIP trunk and SIP proxy servers. Figure 3-5 highlights a deployment of a Cisco PGW VoIP trunking switch in a residential broadband network. It touches aspects of PSTN and IP architecture connectivity. In this particular architecture, the PGW sits at edge of the IP network and deals with offloading the VoIP traffic to the PSTN. The traffic is routed through SIP proxies onto the trunking gateway. Figure 3-5 represents a mixture of networks with various integration boundaries. This shows the SIP connectivity from various sources: the business access, PSTN incoming and outgoing calls, and residential access. All the services provided by these networks need to be tracked. They can be tracked by signaling and media metrics and can help in sizing and service-level assurance (SLA) for the integration points. The SIP, MGCP, and SS7 protocol-related metrics are some of the key metrics that need to be tracked in the IP trunk deployment architecture.

The other major use of IP trunks is across international boundaries, where H.323 networks are prevalent, as seen in Figure 3-6. Figure 3-6 also shows trunk connectivity between the two PGW gateways that are respectively part of large, complex networks. Note the various integration points and the diverse protocol networks. To manage the VoIP service, the capacity and the SLA across these network integration points also become complex. The signaling and other key metrics can help in tracking, trending, and isolating service issues and better plan for capacity and manage SLA.

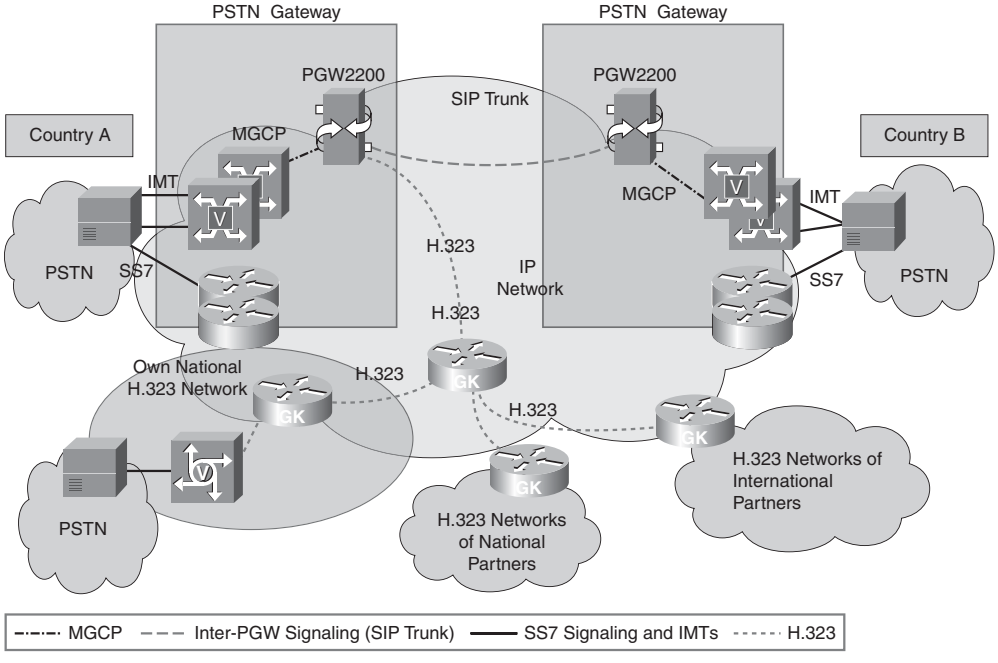


Figure 3-6 *IP Trunk Deployment Across International Boundaries*

In cases in which it is economical to route the traffic over to IP, providers offload the long-distance traffic to another provider rather than using the PSTN. This offloading is provided by a switch that is performing Class IV or toll-switch functionality. You can see in Figure 3-6 that country B connects to country A through an SIP trunk. That way, it can reach H3.323 networks. The PGW keeps track of all CDRs and is extensively used to apply policies for routing traffic through it to optimize cost. In general, services provided by this critical switch need to be tracked. The call and protocol metrics provided by the switch are crucial for running the VoIP network in an efficient way. Thus, the network management capabilities that facilitate this collection and dashboarding become the key to running the VoIP network.

Figure 3-6 shows SIP, MGCP, and H.323 protocols being used for signaling communication. The corresponding traffic counters represent the KPIs needed to effectively monitor the network. The counter collection points are the respective switching, aggregation, and endpoints. Chapter 7 covers these KPIs in detail.

Session Border Controller (SBC) Models

SBCs are also heavily used by VoIP SPs for a variety of reasons. This section discusses some of the common-use cases to continue the discussion of VoIP networks and the corresponding components.

In VoIP networks, the most prevalent and common use for SBCs is offloading VoIP traffic to the PSTN. Other usage includes voice transcoding and network hiding. Figure 3-7 describes a network of an SP with SBC deployed at the edge of the SP network. Figure 3-7 also shows various SBC deployment scenarios. It shows SP1 and SP2 connected to each other through SBCs; this connection is through SIP trunks and is also known as an *SIP tied trunk*. Here, the call flows occur through these SIP trunks from SP1 to SP2. The PSTN connectivity is handled by SP2, representing a PSTN offload network. Another solution for a Call Control Server Farm is also shown in the figure; here, an IMS server is communicating through an SBC to SP1 and at the same time through a set of servers (an SIP Proxy, a Gatekeeper, and MGX/AS5400 trunking gateways) representing the PSTN connectivity network. The other network scenarios reflect the SBC connection from the SBC to an enterprise1 network, a small-business network, and lastly to residential CPE units. All these networks are shown to be carrying different types of traffic along with possibly the VoIP traffic. The key component depicted in the figure is the SBC and how it interconnects with the all the other VoIP networks. The SBC and SIP Router Proxy (SRP) are the tap points for the SIP metrics.

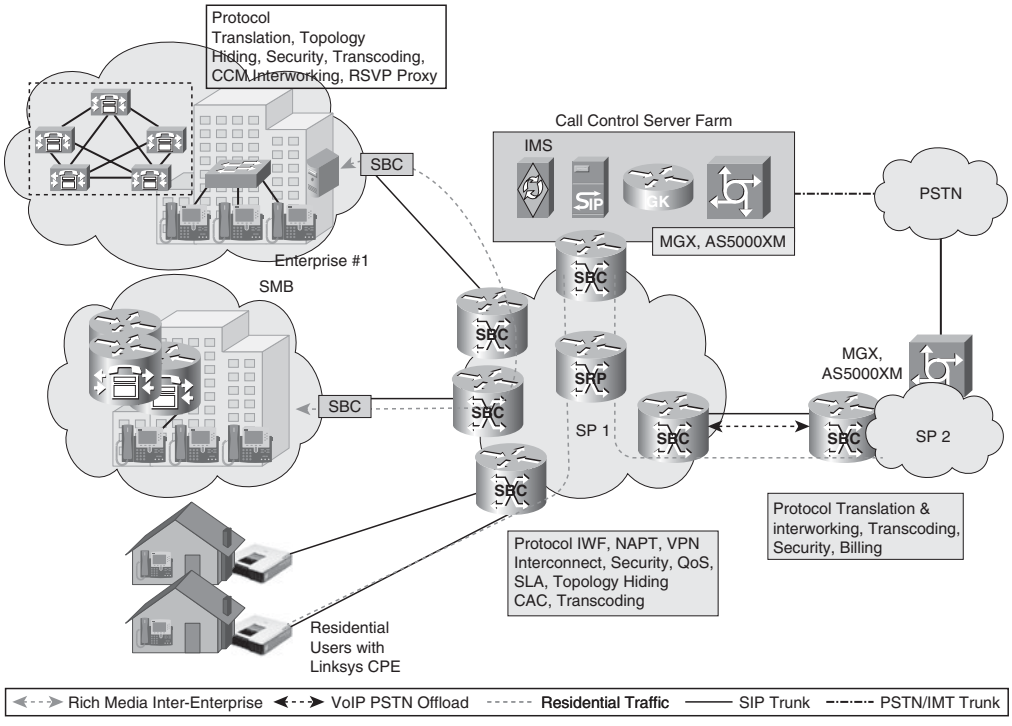


Figure 3-7 *Session Border Controller Deployment Architecture*

Key Components Used in SBC Models

Some of the key components used in SBC deployment models are as follows:

- **SBC:** SBCs are used at the edge of the network; they manage and control the SIP session streams traversing the borders of the networks they sit at. The SBCs provide the functionality for hiding customer networks by NATing. The SIP session streams represent SIP signaling and media communication.
- **SIP Router Proxy (SRP):** SIP Proxy and SRP perform the same functionality. SRPs basically help route requests to the user's current location, authenticate and authorize users for services, implement provider call-routing policies, and provide features to users.

PSTN Offload

VoIP SPs are offloading the handling of PSTN traffic through PSTN SPs. Typically, these are referred to as the traditional telcos. The telco and the VoIP SP have SBC devices at the edges of their networks, and through SIP trunks between these SBCs, the VoIP traffic is offloaded to the PSTN network. The VoIP SP maintains the quality of the VoIP offload service through SLAs. Telcos being the hosts of the VoIP offload also report the SLAs on this service. The VoIP offload allows the VoIP SP to focus and improve the IP-centric traffic and not worry about managing a PSTN-based network.

The needs or challenges for the VoIP SP that drive this design are numerous and include the following:

- Turning up, maintaining, and growing the VoIP network and interfacing with the PSTN involves a lot of overhead.
- The VoIP SP has to plan in advance the trunk capacity needed to service its subscriber base.
- VoIP SPs not only have to deal with the bearer traffic aspect of the service but also the PSTN/SS7 signaling network. Both of these call bearer and signaling aspects have their own infrastructure, and thus are an overhead that needs to be monitored and maintained. The monitoring through select KPIs allows effective operations.
- The PSTN trunk turnup procedure requires countless hours of interaction with the telco; thus Opex cost is high.
- Maintenance is another challenge, especially if operations work is being performed on network augmentation and Emergency/911 circuits. The downtime for this kind of service has many repercussions, some of which can even lead to lawsuits, if the E911 services are impacted. Effective monitoring of the circuits through KPIs allows quick resolution of the outages. The PSTN offload allows the VoIP SP to basically hand over the liability of managing the 911 trunks to the PSTN provider.
- The VoIP SP has to constantly monitor the capacity and continue to profile the traffic to keep up with the subscriber growth.

In addition to the aforementioned needs or challenges, additional aspects include monitoring and managing the complex network of SIP trunks, which are used to interconnect the VoIP switches or the SBCs. Another key challenge is to keep on top of SIP trunk utilization and call performance metrics. As you will see in Chapters 7 and 8, monitoring and correlating these metrics yield an effective VoIP network management system. The SIP traffic counters constitute the KPIs needed to effectively monitor the SIP network. The main theme here is to provide a background of the various operational overheads. Chapters 7 and 8 look at how the metrics are used to track both the hosted PSTN network and the VoIP offload network, which leads to tracking of key metrics.

Network Hiding

Previously, we mentioned that SBCs are deployed at the edge of the VoIP SP network. This section provides more context for this discussion. The SBC enables VoIP signaling and media to be received from and directed to a device behind a firewall and Network Address Translator (NAT) at the border of an adjacent network. The SBC achieves this by rewriting the IP addresses and ports in the call-signaling headers and the Session Description Protocol (SDP) blocks attached to these messages. SDP is basically used for multimedia session setup. This functionality is offered by all SBCs. This NAT functionality enables the VoIP SP provider to hide the network. Hiding allows the SP to not expose its internal network to the outside work, because NAT translates the internal network to another external-facing network. The VoIP SP's SBC basically gets a tied SIP trunk to the SBC of the PSTN provider and does NAT for the back-end internal network.

Being able to look into the traffic enables the SBC to perform a wide range of functionality, including antispam, QoS, and billing. These features also help the VoIP SP to potentially improve the VoIP QoS, better track the connection billing records, and detect any security violations like spam attacks. To summarize, the SBC-based network is effectively monitored thorough KPIs comprised of SIP traffic counters. The collection, correlation, and reporting of these counters are important items that a VoIP SP should perform.

Voice Security in Service Provider Networks

This section discusses some of the security issues related to deploying IP-based telephony services in SP networks. It covers the security challenges seen at the network element and discusses what can be done to address them. Also, signaling and media encryption is discussed to address some of the security issues.

Securing VoIP Network Elements

To prevent DoS attacks and theft of service, the SP can implement security features on network elements that can minimize the chances of an outsider gaining access to valuable network resources and/or free service.

These security measures can include deploying stateful firewalls in the network that allow only authorized traffic to enter the SP network. Configuring access control lists (ACL) on the edge routers can help prevent unwanted traffic in the network. The endpoints that are connected to the edge of the SP network should be authenticated using the Authentication, Authorization, and Accounting (AAA) protocol. Unauthorized users should be denied access to network resources by either black-holing their traffic or assigning them a low bandwidth class of service that would not allow them to send or receive a significant amount of traffic. The key concept behind black-holing is to stop the propagation of this kind of traffic; on the other hand, a low class of traffic has residue in the network.

Securing Call Signaling and the Media

Because call signaling is used for setting up new calls, tearing down calls, and modifying the state of existing calls, it is important that these signaling messages are secured. In the case of a centralized switching model, such as PacketCable, this is accomplished by having a security association between the endpoint and a trusted network device. A security association is a set of provisioned security elements (for example, security keys) on both the endpoint and the trusted network device. By having a set of security associations, the trusted device can authenticate the endpoint when interacting with it. The interaction that takes place between the endpoint and the trusted device can be encrypted. IP Security (IPsec) is one of the mechanisms used to achieve this with the preprovisioned (preshared) keys. This ensures that all traffic between the two devices is from known sources and encrypted.

Similarly, to protect customer privacy, conversations must be kept private. To do this, the media streams generated by customer conversations must be encrypted. The endpoints in the conversation can negotiate a set of ciphersuites (type of authentication and encryption to be used) and then encrypt all their traffic using the negotiated method. Some of the ciphersuites negotiated by the endpoints can include Hash-based Message Authentication Code Message-Digest Algorithm 5 (HMAC-MD5) and Hash-based Message Authentication Code Secure Hash Algorithm (HMAC-SHA) authentication algorithms, and Data Encryption Standard (DES), Triple DES (3DES), and Advanced Encryption Standard (AES) encryption algorithms.

Common Issues and Problems When Deploying IP-Based Telephony Services

This section discusses some of the common issues encountered by providers when deploying IP-based telephony services over their IP infrastructure. The issues are ongoing and need to be maintained for the life of the VoIP service.

Convergence-Related Issues

As discussed earlier in the chapter, VoIP is primarily deployed on converged IP networks. Recall that a converged network is defined as a network capable of transmitting all types of traffic including data, voice, video, and images. Most existing SP IP networks have been designed to carry primarily data traffic and are geared toward data applications such as email, web traffic, and so on. The VoIP traffic is sensitive to time, the packets need to be delivered within a specific time period, and the network needs to facilitate this through various mechanisms. Deploying VoIP in such networks introduces new challenges for the SP's operations staff that needs to carefully monitor the health of their network and work closely with other groups in the company to provide VoIP continuity across the network. For example, in a DOCSIS/IP network, different groups are responsible for managing and maintaining the HFC/RF network, and another group is responsible for the IP network. Although these groups have totally different job responsibilities and technical background, they both need to work closely to provide quality service to the cable providers' customers.

Issues in Media Affecting Quality

Unlike circuit-switched voice networks that have dedicated paths and fixed bandwidth for every call, VoIP networks can share the same resources for data and voice traffic. In some cases, link overutilization or poor media characteristics (noise on RF links) can result in dropped or delayed packets.

VoIP traffic or packetized voice traffic needs to be sent at fixed intervals at the transmitting end so that the receiving end can predictably receive these packets and decode them. Because of serialization delay at the transmitting end, network delay, and jitter, these packets can arrive at the receiving end at varying intervals.

VoIP endpoints use dejitter buffers to compensate for variance in delay during media packet transmission through Real-time Transmission Protocol (RTP). If the dejitter buffers overflow because of excessive delay, this can impact voice quality so that it might sound like a robotized voice. Excessive packet loss can cause issues such as choppy voice quality.

Other voice quality issues can be caused by things such as codec mismatch, where each endpoint uses a different codec (for example, G.711 versus G.729).

Issues in Signaling Affecting the Services and Features

Voice signaling protocols such as MGCP, NCS, and SIP carry important information about how voice calls need to be set up, how resources need to be allocated, and how QoS needs to be provided to the voice traffic.

Network congestion and resource oversubscription can adversely affect voice-signaling protocols that can affect the services and features these protocols support. Voice signaling–related issues can also be caused because of improper network design and misconfigurations.

If voice signaling gets impaired, it can have a range of effects from delayed call setup to failed call setup, from loss of dial tone to one-way voice, and so on.

These issues are often caused by interoperability issues between equipment from different VoIP vendors. Even though they claim to be compliant with protocol standards, they can still have varying implementations of protocol stacks in their products.

IP Routing–Related Issues

SPs might deploy various routing protocols such as Open Shortest Path First (OSPF), Intermediate System–to–Intermediate System (ISIS), Border Gateway Protocol (BGP), and so on for providing IP connectivity across their infrastructure. These routing protocols carry network information that is used for calculating the most efficient path for carrying customer traffic through the SP network.

The failure of these routing protocols can result in a loss of IP connectivity or degraded service for the SP's customers. Such failures can severely impact VoIP traffic. If a link or node in the SP network fails, causing the routing protocol to reconverge or recalculate its routes, the voice traffic might be sent over a low-bandwidth link that can cause voice degradation. For this reason, the SP needs to carefully tweak routing protocol timers to make sure that the network can converge in a timely manner, minimizing the impact to VoIP traffic.

High Availability and Convergence for Business Continuity

A lot of non-Tier1 and non-Tier2 SPs might not implement redundancy in their network when deploying data-only applications. This becomes a critical issue when VoIP is deployed in such networks. A failed router or switch in the SP core network can cause loss of service to data and voice customers.

Therefore, it is critical for the SP to implement redundancy in the network so that a loss of a link or node does not result in loss of service to its users. Implementing redundancy in the network might involve deploying hardware and software with high-availability features. Redundancy is implemented at a device level where the hardware has active and standby components. If the active component fails, the standby can take over without causing a network outage. Redundancy is also implemented at a link level where multiple links provide connectivity to other network resources, so if a link fails, the other links can carry all the traffic. In some cases, the SP might also choose to deploy redundancy in the form of additional hardware that can take over if certain devices in the network fail. Redundancy can also be implemented in software, such as in routing protocol implementations, which can provide alternate routes through the SP network if the primary/best path fails.

The focus here is not the type of redundancy, or the various specific challenges, but the impact of the failures. Thus, effectively tracking key metrics can help in sustaining the VoIP service. These metrics can range from protocol to Layer 2/3 and h/w uptime metrics.

Summary

This chapter described some of the common SP voice deployment models that are deployed by major SPs in the United States and around the world. These networks are evolving, and a mind-set needs to be created for identifying a key performance indicator (KPI) across these networks. Thinking at the protocol layers and identifying network elements that can be collection points for these protocols are the key. After these KPIs are identified, monitoring them through a series of dashboards allows the SP to effectively manage its services.

References

1. Riddel, Jeff. *PacketCable Implementation*. Indianapolis, IN: Cisco Press, 2007.

2. Davis, Brian. *PacketCable Primer*. Cisco Systems, Inc.

Index

A

B

C

D

E

F

G-H

I

J-K

L

M

N

O-P

Q-R

S

T

U

V

W

X-Y-Z