

Matthew Stafford

Signaling
and **Switching**
for Packet Telephony

TN916.2
S779

Signaling and Switching for Packet Telephony

Matthew Stafford



Artech House, Inc.
Boston • London
www.artechhouse.com

Library of Congress Cataloguing-in-Publication Data

Stafford, Matthew.

Signaling and switching for packet telephony / Matthew Stafford

p. cm.—(Artech House telecommunications library)

Includes bibliographical references and index.

ISBN 1-58053-736-7 (alk. paper)

1. Internet telephony. 2. Packet switching (Data transmission) I. Title. II. Series.

TK5105.8865.S73 2004

621.382'16—dc22

CIP

2004053829

British Library Cataloguing in Publication Data

Stafford, Matthew

Signaling and switching for packet telephony. —(Artech House telecommunications library)

1. Internet telephony 2. Packet switching (Data transmission)

I. Title

621.3'85

ISBN 1-58053-736-7

Cover design by Yekaterina Ratner

© 2004 Matthew Stafford and Cingular Wireless. All rights reserved.

Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House 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.

Library of Congress Cataloguing-in-Publication Number: 2004053829

International Standard Book Number: 1-58053-736-7

10 9 8 7 6 5 4 3 2 1

Signaling and Switching for Packet Telephony

Acknowledgments

During the development of this book, I had many helpful conversations with colleagues. I would like to thank the following people in particular: Haifeng Bi, Mike Boeckman, Jasminka Dizdarevic, Steve Frew, Cathleen Harrington, David Howarth, Rich Kobylinski, Si-Jian Lin, Rias Muhamed, Jessie Lee, Steve Partridge, Simon Richardson, Sam Sambasivan, Richard Tam, Randy Wohlert, and Mark Wuthnow.

The preparation of this book devoured many evenings and weekends. I would like to thank my wife, Miriam, for her patience.

Contents

Acknowledgments	<i>xiii</i>
CHAPTER 1	
Introduction	1
1.1 In the Beginning, There was Voice	1
1.2 Motivation: What Is the Case for Packet Telephony?	2
1.2.1 One Network Versus Two	2
1.2.2 Services	3
1.3 Switch Design	3
1.3.1 Separating Bearer and Control Planes	4
1.4 Motive and Opportunity for Carriers	5
1.5 What Are We Waiting For?	6
1.6 Motivation for this Book	7
PART I	
Switching Architectures for Packet Telephony: An Expository Description	9
CHAPTER 2	
Essentials of Next Generation Switching	11
2.1 Another Look at the Backhaul Example	12
2.2 Ability to Enter New Markets	13
2.3 Switch Components and Terminology	14
2.3.1 Where Does One Switch Component End and Another Component Begin?	14
2.4 A Useful Abstraction	15
2.5 Defining the Fabric	16
2.5.1 Do Control Messages Between Media Gateways and Their Controller Pass Through the Switch Fabric?	17
2.5.2 What Is a Packet?	18
CHAPTER 3	
Motivation for Packet Telephony Revisited	21
3.1 Separation of Bearer and Control	21
3.1.1 Open Interfaces	22
3.1.2 Introducing and Maintaining Services	23

3.1.3	New Bearer Types	25
3.2	Packet Fabrics	26
3.2.1	Exploiting Routing Intelligence of Packet Networks	26
3.2.2	Exploiting Low Bit-Rate Voice Codecs	29
CHAPTER 4		
	Signaling and Services	31
4.1	The Control Plane	31
4.2	What Is a Service?	32
4.2.1	Vertical Services	32
4.2.2	Services that Offer Alternative Billing Schemes	33
4.2.3	Short Message Service	33
4.3	Where Do Services “Live,” and What Do They Entail?	33
4.3.1	Can You Say “Database?”	34
4.4	Limitations of Circuit-Switched Networks	35
PART II		
	Components of Packet Telephony: Technical Descriptions	37
CHAPTER 5		
	Introduction to Part II	39
5.1	Selected Telco Terminology	40
CHAPTER 6		
	Protocols	43
6.1	What Is a Protocol Stack?	43
6.1.1	Comparison with Last In, First Out Data Structures	44
6.2	Generic Layer Descriptions	44
6.2.1	Data Link Layer	45
6.2.2	Network Layer	46
6.2.3	Transport Layer	46
6.2.4	A Note on Terminology: Packets and Frames	46
6.2.5	General Comments	47
6.3	Internet Protocol and Transmission Control Protocol	48
6.3.1	What Is an Internet Protocol Router?	48
6.3.2	A Brief Look at TCP	48
6.3.3	TCP/IP Networking Illustration	51
6.3.4	Alternatives to TCP at Level 4: UDP and SCTP	52
6.4	What Is a Finite State Machine?	53
6.4.1	States	54
6.4.2	State Transitions	54
6.4.3	Additional Comments	55
6.5	Signaling System 7 in Brief	55
6.5.1	MTP2	56
6.5.2	MTP3	57
6.5.3	SCCP	57
6.5.4	TCAP	57

6.5.5	MAP	57
6.5.6	ISUP	58
6.6	Summary	60
	References	61

CHAPTER 7

A Closer Look at Internet Protocol		63
7.1	The IPv4 Header	64
7.1.1	Fragmentation and Path MTU Discovery	65
7.2	The IPv6 Header	65
7.2.1	IPv6 Extension Headers	66
7.3	Addressing and Address Resolution	67
7.3.1	Conserving IPv4 Address Space	67
7.3.2	The IPv6 Address Space	68
7.3.3	Uniform Resource Identifiers and Domain Name System	69
7.4	Security and AAA	69
7.4.1	Security	69
7.4.2	Authentication, Authorization, and Accounting	70
7.5	Routing	71
7.5.1	Network Optimization	71
7.5.2	Internet Routing Protocols	72
7.5.3	A Link State Protocol: OSPF	72
7.5.4	Distance Vector Protocols: RIP and BGP	73
7.5.5	Routing Protocol Convergence	73
7.5.6	Scalability	75
7.5.7	Trade-offs	75
7.6	Reachability Information	76
7.7	Quality of Service and Statistical Multiplexing	76
7.7.1	What Is Statistical Multiplexing?	77
7.7.2	Differentiated Services	78
7.7.3	Multiprotocol Label Switching	79
7.7.4	“DiffServ at the Edge, MPLS in the Core”	80
7.7.5	Multiservice Networks	80
7.8	Layer 4 Protocols: Suitability to Task	81
7.8.1	UDP	81
7.8.2	Carrying SS7 Traffic over an IP Network: SCTP	82
7.8.3	Comparing and Contrasting TCP with UDP and SCTP	83
7.9	Mobile IP	84
7.10	Summary	84
7.10.1	Further Reading	85
	References	85

CHAPTER 8

A Closer Look at SS7		89
8.1	SS7 Architecture and Link Types	89
8.2	SS7 Routing and Addressing	91
8.3	Review of the SS7 Protocol Stack	92

8.4	Message Transfer Part	93
8.4.1	MTP2	94
8.4.2	MTP3	94
8.5	SCCP	95
8.5.1	General Description and Communication with MTP3	95
8.5.2	Getting There Is Half the Fun: Global Title Translation	96
8.6	TCAP	98
8.6.1	Number Portability	99
8.7	MAP	100
8.8	Summing Up	103
8.8.1	Additional Weaknesses of SS7	104
8.8.2	Strengths of SS7	105
	References	105
CHAPTER 9		
	The Bearer Plane	107
9.1	Voice Encoding	107
9.1.1	G.711	107
9.1.2	Why Digital?	108
9.1.3	Other Voice-Encoding Schemes	108
9.2	Bearer Interworking	111
9.2.1	Transcoding	111
9.2.2	Encapsulation of Digitized Sound	111
9.2.3	Packetization Delay and Playout Buffers	113
9.3	Voice over IP	113
9.3.1	Real-Time Services in IP Networks: RTP over UDP	113
	References	116
CHAPTER 10		
	Media Gateway Control and Other Softswitch Topics	119
10.1	Requirements	119
10.1.1	ID Bindings	121
10.2	SDP in Brief	122
10.3	Megaco/H.248	123
10.3.1	Introducing the Megaco Connection Model	123
10.3.2	Terminations	124
10.3.3	Contexts	124
10.3.4	Megaco Commands	125
10.3.5	Example Call Flow	126
10.3.6	Usage of the Move Command	129
10.3.7	Descriptors	130
10.3.8	Sample Megaco Messages	132
10.3.9	Three Way-Calling Example	134
10.3.10	Megaco Miscellanea	136
10.4	MGCP	137
10.4.1	Example Call Flow	137
10.4.2	Brief Comparison with Megaco	138

10.4.3	Other MGCP Verbs	140
10.4.4	Transactions and Provisional Responses	141
10.4.5	MGCP Packages	142
10.5	Interworking with Circuit-Switched Networks	142
10.5.1	Latency Trade-offs	142
10.6	Inhabiting the Bearer, Service, and Control Planes	143
10.7	Signaling Between Two Softswitches	143
10.7.1	BICC	143
	References	144

CHAPTER 11

	Session Control	145
11.1	“Generic” Session Control	145
11.1.1	Comparison with ISUP Call Flow	147
11.1.2	Modularity in Protocol Design	147
11.2	The H.323 Protocol Suite	148
11.2.1	Heritage of H.323: ISDN	148
11.2.2	H.323 Call Control and Media Control Signaling	149
11.2.3	Talking to the Gatekeeper: RAS Signaling	150
11.2.4	Evolution of H.323	151
11.3	SIP Basics	152
11.3.1	SIP Requests and Responses	154
11.4	SIP Functional Entities	155
11.4.1	Proxy Servers and Redirect Servers	156
11.4.2	Back-to-Back User Agents	157
11.4.3	Registrars	157
	References	157

CHAPTER 12

	More on SIP and SDP	159
12.1	A Detailed SDP Example	159
12.1.1	Additional Line Types	161
12.2	A Detailed SIP Example	161
12.2.1	Registration Procedures	161
12.2.2	Making a Call	164
12.2.3	The Offer/Answer Model	167
12.3	Forking of SIP Requests	168
12.4	SIP for Interswitch Signaling	168
12.4.1	Comparison with BICC	170
12.5	Additional SIP Methods	170
12.5.1	UPDATEs and re-INVITEs	171
12.6	Resource Reservation and SIP	171
12.6.1	QoS Attributes in SDP	173
12.6.2	More on Parameter Negotiation	174
12.7	SIP-T and Beyond	174
12.8	Authentication and Security	176
12.9	Further Reading	176

References	177
CHAPTER 13	
Implementing Services	179
13.1 Introduction	179
13.2 SS7 Service Architectures: Intelligent Network	180
13.2.1 The Global Functional Plane	182
13.2.2 The Distributed Functional Plane	182
13.2.3 IN Capability Sets	183
13.2.4 Limitations and Trade-offs of IN	184
13.3 CAMEL and WIN	185
13.4 Parlay/OSA	185
13.5 JAIN	186
13.6 SIP and Services	186
13.6.1 SIP and Intelligent Networks: PINT and SPIRITS	186
13.7 SIP in Wireless Networks	190
13.7.1 Push To Talk over Cellular	190
13.7.2 SIP Header Compression	191
13.7.3 IP Multimedia Subsystem	192
13.8 Short Message Service	195
13.8.1 SMS in Support of Other Applications	196
13.9 Further Reading	196
References	197
CHAPTER 14	
Properties of Circuit-Switched Networks	199
14.1 Telco Routing and Traffic Engineering	199
14.1.1 Truitt's Model	200
14.1.2 Dynamic Nonhierarchical Routing, Metastable States, and Trunk Reservation	202
14.1.3 Optional Section: Traffic Intensity and the Erlang B Formula	203
14.2 Comparison with IP Routing and Dimensioning	205
14.3 Security	206
14.4 Quality of Service	206
14.5 Scalability	207
14.6 Survivability and Reliability	207
14.7 Billing Functionality	207
14.8 Emergency Service and other Government Mandates	208
References	208
CHAPTER 15	
Evolving Toward Carrier-Grade Packet Voice: Recent and Ongoing Developments	209
15.1 QoS and Traffic Engineering in IP Networks	209
15.1.1 Class-Based Queuing	209
15.1.2 DiffServ and IntServ Revisited	210

15.1.3	Verifying and Enforcing Traffic Contracts	212
15.1.4	ITU-T and 3GPP QoS Standards	212
15.2	Service-Level Agreements and Policy Control	215
15.3	SDP and SDPng	216
15.4	Sigtran Adaptation Layers	216
15.5	Middlebox Traversal	217
15.6	Comments and Further Reading	219
15.6.1	More on IP QoS	219
15.6.2	IPv6 and ROHC	219
15.6.3	Routing for Voice over IP Protocols: iptel Working Group	221
15.6.4	ENUM	221
15.6.5	Service Architectures	222
	References	222

CHAPTER 16

Conclusion	225
------------	-----

APPENDIX A

Data Link Layer Protocols	227
A.1 HDLC	227
A.2 Frame Relay	227
A.2.1 The Frame Relay Header	228
A.2.2 Label Switching and Virtual Circuits	229
A.3 Asynchronous Transfer Mode	230
A.3.1 The ATM Header	230
A.3.2 ATM Approach to Quality of Service and Statistical Multiplexing	230
A.3.3 The ATM Control Plane	231
A.3.4 ATM Adaptation Layers and Options for Voice over ATM	232
A.3.5 Virtual Paths	234
A.3.6 MPLS over ATM: VC Merge Capability	234
A.3.7 Why Not Voice over ATM?	235
A.4 Ethernet	236
A.4.1 History of Ethernet	237
A.4.2 Ethernet Frame Structure	237
A.4.3 CSMA/CD and Its Scalability Limitations	238
A.4.4 Hubs, Bridges, and Switches	238
A.4.5 Further Reading	240
References	240

About the Author	243
Index	245

Introduction

1.1 In the Beginning, There was Voice

Voice telephony was arguably the first telecommunications service to achieve truly widespread deployment. Data services came into their own much later. To be sure, data telecommunication (in the form of smoke signals and other visual semaphore schemes, for example) has been around for a very long time. Note also that the telegraph preceded the telephone, and that innumerable messaging schemes of all sorts have been devised and used for military purposes. Morse code telegraphy is an interesting example of a data service: this scheme for electrical transmission of text messages was commercially viable before telephony became available.

In scale of deployment, however, telegraph service never came close to the level subsequently reached by telephony. The average person in a developed country has direct access to telephones at home and at work. Moreover, this has been the case for many years; nowadays, wireless telephones add mobility to the equation.

Data telecommunication became common in the consumer market only within the last 10 to 20 years. Uptake in the academic and business communities came somewhat earlier (but still much later than voice). Thus voice networks were already ubiquitous when data networking technologies reached mass-market scale, and it was perfectly natural to ask whether one could also use these networks to transmit large volumes of data between distant sites.

In a nutshell, the question was:

“We have extensive voice networks; are they useful for carrying data as well?”

The answer to this question was definitely “Yes!” The biggest reason for this resounding affirmative was that, starting in the 1960s, the voice network in the United States evolved from analog to digital. In other words, this network was already transporting bits.

For years, data networks were islands in a voice-based world. Data networking technologies designed for so-called local area networks (LANs) suffered from severe distance limitations. Long-haul data transmission in the consumer, academic and business markets was achieved using networks initially designed for voice. In fact, telephone carriers’ networks provided the only viable means of interconnecting distant data networks. In many instances, data traffic is still transported in this fashion; fax and dial-up modem transmissions are familiar examples. Widespread access to dedicated wide area data networks (such as the Internet) is a relatively recent phenomenon.

Dedicated data networks are increasingly common and far-reaching, however. This begs the following question: “We have extensive data networks; are they useful for carrying voice traffic?” (Although voice telephony is our primary focus in this book, note that the question remains equally pertinent for other real-time services such as videoconferencing.)

Today, telephone networks and data networks are very different beasts, designed according to different philosophies. Here is a crucial distinction to keep in mind:

- In traditional telephone networks, transmission capacity is meted out in a continuous fashion. That is, a fixed amount of capacity is allocated to each call and deallocated when the call ends. Meanwhile, this transmission capacity cannot be shared with any other call (even when the parties on the first call are silent). We say that traditional telephone networks are circuit-switched; here the term circuit refers to the end-to-end transmission capacity that is reserved throughout the life of a call.
- In data networks, transmission capacity is allocated in a discrete fashion. Suppose two pairs of users are conducting ongoing sessions that use a shared transmission link. When transmission of a packet for one session is completed, the other session can transmit a packet, although neither session has ended. Here a packet is a chunk of digital data (i.e., a sequence of bits); we say that data networks are packet-switched.

Note that these are somewhat oversimplified in the interest of brevity.

1.2 Motivation: What Is the Case for Packet Telephony?

1.2.1 One Network Versus Two

To understand why there is a great deal of interest in packet telephony, one need look no further than the corporate environment. A typical office site has two completely separate in-building networks: a circuit-switched network for telephones and a packet-switched network for computers. The latter is called a local area network. Each of these networks must be provisioned and maintained. Could we combine telephone and data traffic on one of the two networks (thereby allowing us to reduce cost by eliminating the other network)?

Of the two in-building networks, the computer network undoubtedly has much greater bandwidth than the phone network. Thus, if we try to interconnect our computers via the phone lines (abandoning the local area network), there is essentially no hope of a satisfactory result.

Therefore, if we want to dispense with one of the two networks, eliminating the in-building phone network is the only reasonable choice. And this option entails placing voice traffic over a packet-based medium.

The situation is substantially different in the consumer market. A local telephone company that already has a revenue-producing “legacy” voice network has little motivation to invest in voice over packet technologies. However, cable

companies might become interested in doing just that, so as to compete with local telephone companies. In recent years, cable companies have upgraded their networks so they can provide broadband Internet access. In the process, cable has become a bidirectional packet-based medium, suitable for packet voice traffic. We note that, as of this writing, cable telephony has not taken off (at least in the United States).

1.2.2 Services

Packet telephony solutions must provide some economic benefit (that is, increased revenue and/or reduced cost); otherwise, they will not be widely deployed. We have already begun to address reduced cost with our previous example. We will return to this topic later, exploring the service provider's point of view.

Now we look at revenue generation. To increase revenue, a telephone service provider must add new customers (and this is becoming increasingly difficult) or must come up with new features that customers will want to buy. This brings us to the development of enhanced services.

As an example, we look at "find me/follow me" services. The basic idea of find me/follow me is a simple one: to offer flexible configuration options for call-forwarding behavior on a per-user basis. Suppose, for example, a traveler wants certain callers (identified by calling party numbers, say) to be able to reach him/her via automatic forwarding from his/her landline phone to a wireless phone. All other calls will be forwarded to voice mail. Such features are offered by today's Private Branch Exchanges, but are not generally available to consumers.

Moreover, it is easy to envision useful "add-on" functionalities that are difficult to implement in today's networks. For example, it would be nice if one could configure out-of-office settings for voice mail and e-mail from the same menu, perhaps employing a speech-to-text processor to convert the voice mail greeting to a text message that is automatically sent in reply to incoming e-mail messages.

Other desirable options include the ability to reconfigure (wireline) forwarding options from a wireless phone. This is difficult partly because wireless and wireline networks grew up separately; so-called "Intelligent Network" features in the two realms are based on different signaling protocols. (We flesh out this topic in Chapter 13.) Moreover, with the User Interface limitations of today's phones, subscribers may find such features difficult to use. Admittedly, packetized voice does not automatically bring about "convergence" between wireless and wireline networks. Well-designed service control schemes that can cross network boundaries, however, may facilitate convergence. Therefore such schemes promise to be an important part of the overall evolution toward packet telephony.

1.3 Switch Design

Packet voice equipment is available and in use today, so packet-based telephony is certainly feasible. If people only wanted to call others located in the same building as themselves, the case for packet telephony in the corporate environment would be overwhelmingly positive.

In reality, one of the best features of existing telephone networks is that it is possible to call almost anyone. To maintain this universality, interworking with the outside world is a must. This boils down to interworking with circuit-switched networks, since the vast majority of telephones today are connected to circuit switches. In particular, circuit switching is utterly predominant in public telephone networks; packet voice is only just beginning to make inroads in this market.

How should packet voice switches be designed? This is one of the main topics of discussion in this book. Specifically, we will talk extensively about the telco environment and design principles that are expedient for operating in this environment. In this context, we will draw comparisons with legacy voice switches. We note that the design principles discussed here are equally applicable to cable operators, if and when they decide that they want to become large-scale telephone service providers.

1.3.1 Separating Bearer and Control Planes

The separation of bearer and control planes is a fundamental concept in next generation switch design. The bearer plane is the part of the network that carries end-user traffic (e.g., voice samples, in the case of telephony). As the name suggests, the control plane is the part of the network that carries call-control signaling. In circuit switches, the bearer and control planes are not clearly separated.

In a nutshell, the reason for separating the bearer and control planes is the promise of increased flexibility. This flexibility can take several forms, notably:

- *Distributed architecture.* A rich set of options for placement of switch components: the elements of a switch can be geographically dispersed.
- A rich set of options [based on packet technologies such as Internet Protocol (IP), Asynchronous Transfer Mode (ATM) and Ethernet as well as sophisticated vocoders] for representing, encapsulating, routing, and transporting voice traffic.
- The ability to base the creation and implementation of new services on standardized open interfaces. This is an important step toward the “holy grail” of services that combine voice, video, and data in useful ways.
- Flexibility in choosing suppliers. That is, different components can potentially be purchased from different equipment vendors.

We will return to these topics in the Section 1.4 (where we argue that these advantages make a compelling case for packet telephony) and elsewhere.

At this point, the reader can begin to see that packet telephony is much more than replacing circuit-switched bearer channels with packet-switched alternatives. It is possible to build switches that internally employ packet bearers, but for all intents and purposes act exactly like circuit switches. There is a place for such technology. However, we will see that next generation switching concepts hold the potential for much more.

1.4 Motive and Opportunity for Carriers

Why would a telephone service provider want to invest in voice over packet technology? We saw one reason in Section 1.2: to enable enhanced services (and thereby realize new revenue streams). In the following example, the goal is to reduce cost.

We said earlier that separation of bearer and control allows for the components of a switch to be geographically dispersed. To see why this is important, we direct the reader's attention to Figure 1.1. (The *fabric* of a switch is the conduit through which voice samples flow from the calling party to the called party and vice versa. Each *area* in the figure might represent a local switch, along with all of the customers that are homed to that switch, or a private branch exchange, etc.) Note that areas 1 and 2 are not directly connected. Therefore, when a customer in area 1 calls a customer in area 2, the bearer path must include the nearest switch that connects to both areas (switch A in this example). If areas 1 and 2 are much closer to one another than they are to switch A, then we are faced with so-called “backhaul” costs. That is, the voice-encoding bit stream must travel the long way around for the duration of the call.

Suppose that the volume of traffic between areas 1 and 2 is not sufficient to justify either of the following alternatives:

- Reserving dedicated transmission capacity between areas 1 and 2;
- Installing an additional switch closer to areas 1 and 2.

Suppose, however, that the volume of traffic is large enough that it grieves us to pay for backhaul transmission capacity. If, by shifting to a distributed design, we could dramatically reduce the cost of adding switching capacity at a nearby location, then we would have a viable alternative.

In Figure 1.2, we illustrate the notion that the fabric under the command of a given controller can consist of geographically-dispersed nodes. If these fabric nodes are very inexpensive (relative to the cost of a “legacy” voice switch), and if one controller can direct the operation of many fabric nodes, then the cost of introducing switching capacity to new locations can indeed be reduced a great deal.

Note that a new type of traffic appears in Figure 1.2: control messages between the controller and the fabric component at location B. (The Controller also sends commands to the colocated fabric component at location A, but these messages do not require interlocation transmission facilities.)

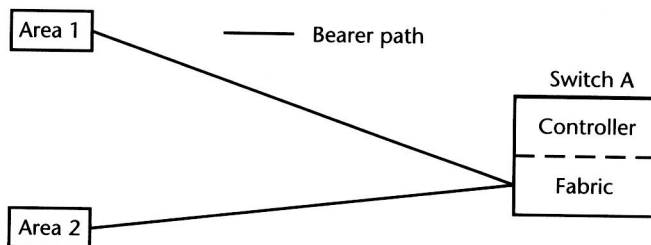


Figure 1.1 Bearer path must traverse closest switch.