SIP

Understanding the Session Initiation Protocol

Third Edition

Alan B. Johnston

SIP: Understanding the Session Initiation Protocol

Third Edition

Alan B. Johnston

ARTECH
HOUSE

BOSTON | LONDON
artechhouse.com

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the U.S. Library of Congress.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Cover design by Yekaterina Ratner
Cover art by Lisa Johnston

ISBN 13: 978-1-60783-995-8

685 Canton Street
Norwood, MA 02062

10 9 8 7 6 5 4 3 2 1

SIP: Understanding the Session Initiation Protocol

Third Edition

For a complete listing of titles in the
Artech House Telecommunications Series,
turn to the back of this book.

For Lisa

Foreword to the First Edition

The Internet now challenges the close to $1 trillion world telecom industry. A renaissance in communications is taking place on the Internet. At its source are new communication protocols that would be impractical on the centralized control systems of circuit-switched networks used in telecommunications. The Internet and the World Wide Web can be technically defined only by their protocols. Similarly, IP telephony and the wider family of IP communications are defined by several key protocols, most notably by the Session Initiation Protocol, or SIP.

The previously closed door of telecommunications is now wide open to web developers because of SIP and its relation to the web HTTP 1.1 protocol and the e-mail SMTP protocol. IP communications include voice/video, presence, instant messaging, mobility, conferencing, and even games. We believe many other communication areas are yet to be invented. The integration of all types of communications on the Internet may represent the next "killer application" and generate yet another wave of Internet growth.

As explained in this book, SIP is a close relative of the HTTP 1.1 and SMTP protocols. This represents a revolution in communications because it abandons the telecom signaling and control models developed for telephony over many years in favor of Internet and web-based protocols. Users and service providers obtain not only seamless integration of telephony and conferencing with many other World Wide Web and messaging applications, but also benefit from new forms of communications, such as presence and instant messaging.

Mobility can also be managed across various networks and devices using SIP. Location management is now under user control, so that incoming "calls" can be routed to any network and device that the called party may prefer. Users may even move across the globe to another service provider and maintain not only their URL "number", but also their personal tailored services and preferences. The end user gains control over all possible preferences, depending on various parameters such as who the other party is, what network he is on and what devices he is using, as well as time of day, subject, and other variables.

The new dimension in communications called "presence" enables users for the first time to indulge in "polite calling" by first sensing presence and preferences of the other party, before making a call. In its turn, presence can trigger location- and time-dependent user preferences. Users may want to be contacted in different ways, depending on their location and type of network access.

E-commerce will also benefit from IP communications. Extremely complex telecom applications, as found in call centers, have become even more complex when integrated with e-mail and web applications for e-commerce. Such applications, however, are quite straightforward to implement using SIP, due to its common structure with the web and e-mail. For example, both call routing and e-mail routing to agents—based on various criteria such as queue length, skill set, time of day, customer ID, the web page the customer is looking at, and customer history—can be reduced to simple XML scripts when using SIP and another IETF standard, the Call Processing Language (CPL). These examples are in no way exhaustive, but are mentioned here as a way of introduction.

This book starts with a short summary of the Internet, the World Wide Web, and its core protocols and addressing. Though familiar to many readers, these chapters provide useful focus on issues for the topics ahead. The introduction to SIP is made easy and understandable by examples that illustrate the protocol architecture and message details. Finally, in the core of the book, a methodical and complete explanation of SIP is provided. We refer the reader to the Table of Contents for a better overview and navigation through the topics.

Alan Johnston has made significant contributions toward the use of SIP for communications over the Internet. I had the privilege of watching Alan in meetings with some of the largest telecom vendors as he went methodically line by line over hundreds of call flows, which were then submitted as an Internet Draft to the Internet Engineering Task Force (IETF) and implemented in commercial systems. Alan combines in this book his expertise and methodical approach with page turning narrative and a discreet sense of humor.

I could not help reading the book manuscript page by page, since everything from Internet basics, protocols, and SIP itself is explained so well, in an attractive and concise manner.

Henry Sinnreich
Distinguished Member of Engineering
WorldCom
Richardson, Texas
July 2000

Preface to the Third Edition

Like the SIP protocol, this book continues to expand and grow in new directions. While the core SIP protocol is essentially the same as it was in 2000, in the years since then, many SIP extensions have been proposed and standardized in the IETF. Important areas such as NAT traversal, SIP and media security, and peer-to-peer (P2P) applications have made many changes in the protocol and the protocol suite. As before, this book will help you sort through all the RFCs and fully understand this important protocol.

Chapter 1 is an introduction for those unfamiliar with the basics of the Internet and the TCP/IP protocol suite. Readers with familiarity in this area can probably skip straight to Chapter 2 for a quick introduction to the SIP protocol. For the rest of the readers, you will find an overview of the various layers of the protocol stack and an introduction to such concepts as the Domain Name Service (DNS) and Uniform Resource Locators (URLs) and an introduction to the Internet standards process in the Internet Engineering Task Force (IETF). Also for the beginner, a new appendix introduces the basics of ABNF and XML, essential for understanding SIP syntax.

Chapters 2 and 3 give you the basics of the protocol, the various messages, elements, transports, and applications. Chapters 4 and 5 explain the SIP request types or methods, with one new method introduced in this edition. Chapters 5 and 6 detail all the SIP response messages and header field types. There are literally dozens and dozens of new responses and header fields in this edition; Chapter 6 alone now has 50 references, five times as many as the first edition. Chapter 7 covers mobility and wireless aspects of SIP. Chapter 8 covers the important presence and instant messaging applications and extensions for SIP, known as SIMPLE. While the basics for these applications were in previous editions, the full suite of functions and operations is now standardized and covered with examples in this chapter. Chapter 9 covers SIP services from VoIP to SIP trunking to conferencing, fax, and video.

Chapter 10 represents many years of hard work in the industry for solving the Network Address Translation (NAT) traversal issue. A large impediment to SIP and VoIP deployment has been NAT traversal, and the scalability of many older solutions has been a major headache for deployments. This chapter explains how NAT works, new classifications schemes for understanding the types of NAT. Then, hole punching is explained, with detailed examples of various types of NAT. NAT traversal protocols such as STUN, TURN, and ICE are explained, and the ways that SIP uses them to scaleably solve the NAT traversal problem are described. This chapter represents the latest technology from the field, presented in this edition for the first time.

This edition also has more information about related protocols such as SDP, H.323, Jabber or XMPP, and Real-Time Transport Protocol (RTP) for media. Chapter 13 discusses how SIP negotiates various types of multimedia sessions.

Chapter 14 is a concise summary of SIP security, starting with the basics of security and encryption, attacks, security protocols, and SIP authentication mechanisms. This chapter also has detailed information about delivering secure media (Secure RTP or SRTP) with SIP. New media keying protocols including ZRTP are also covered.

Chapter 15 is new, covering peer-to-peer (P2P) technologies and how they apply to SIP. The new RELOAD protocol being developed in the IETF is covered, along with other approaches including Host Identity Protocol or HIP.

Chapter 16 includes the detailed call flows for which this book is famous, and Chapter 17 discusses the future of SIP and areas of development and innovation in the years to come. Perhaps these topics will find their way into another future edition.

Also new in this edition are the review questions at the end of most chapters. This has evolved from the classes I have taught on Internet communications at Washington University in St. Louis over the past few years. In fact, much of the new material was generated and developed for my teaching. As a result, I thank my past (and future) students for their interest and attention. I also thank Dr. John Corrigan and Dr. Tom Bush for their support and encouragement of this class. I would like to thank Anwar Siddiqui, Harvey Waxman, and Mun Yuen Leong at Avaya for their support of my SIP work. I especially thank my wife Lisa for her excellent cover artwork and figure design.

In closing, I thank everyone at Artech for giving me the opportunity once again to write about my favorite topic. I also thank you, my past and current readers, for your interest, support, and enthusiasm, and I wish you the best of luck in all your endeavors with the Session Initiation Protocol.

Preface to the Second Edition

Much has changed in the 2.5 years since the first edition of *SIP: Understanding the Session Initiation Protocol* was published. In 2001, SIP was a relatively unknown quantity, an upstart in the voice over IP (VoIP) and multimedia communications industry. Today, SIP is seen as the future of call signaling and telephony. It has been widely deployed by service providers and enterprises and is used casually every day by users of the dominant PC operating system. The full range of possibilities enabled by SIP is just now being glimpsed, and many more possibilities are yet to come.

One reason for this rapid acceptance is that SIP is an incredibly powerful call control protocol. It allows intelligent end points to implement the entire suite of telephony, Private Branch Exchange (PBX), Class, and Centrex services without a service provider, and without a controller or switch, for example.

The biggest driver for SIP on the Internet, however, has less to do with SIP's signaling and call control capabilities. Instead, it is due to the extensions of SIP that turn it into a powerful "rendezvous" protocol that leverages mobility and presence to allow users to communicate using different devices, modes, and services anywhere they are connected to the Internet. SIP applications provide support for presence—the ability to find out the status or location of a user without attempting to set up a session.

Another major change in the past few years is the adoption of wireless SIP to enable multimedia IP communications. As described in the chapter on wireless, SIP is now being used both in its standard form over 802.11 wireless networks and in planned commercial Third Generation Partnership Project (3GPP) rollouts in the coming years. SIP is ideally suited for this key application.

Since 2001 SIP has also grown in terms of the specification itself. Initially, SIP was described by a single RFC with a few related RFCs and a couple of RFC extensions. In Chapter 6 alone, more than 20 SIP-related RFCs are referenced. This book attempts to put all those documents together and provide a single reference for the protocol and all its extensions. Even SIP headers and responses that were standardized in the past but are now removed (deprecated) are listed in this text, providing useful context and background. Many others are discussed that are in the final stages of standardization prior to publication as RFCs, providing an up-to-date insider's view of the future of the protocol. In closing, I again thank my colleagues in the Internet Engineering Task Force (IETF) and at MCI for all their contributions to the development of this protocol—it has been a privilege to be a part of a group of people that have created the SIP industry. Finally, I'd like to tip my hat to two of the key inventors of SIP who continue to develop and propel its implementation: Henning Schulzrinne and Jonathan Rosenberg.

Preface to the First Edition

When I began looking into the Session Initiation Protocol (SIP) in October 1998, I had prepared a list of a half dozen protocols relating to Voice over IP and Next Generation Networking. It was only a few days into my study that my list narrowed to just one: SIP. My background was in telecommunications, so I was familiar with the complex suite of protocols used for signaling in the Public Switched Telephone Network. It was readily apparent to me that SIP would be revolutionary in the telecommunications industry. Only a few weeks later I remember describing SIP to a colleague as the "SS7 of future telephony"—quite a bold statement for a protocol that almost no one had heard of, and that was not even yet a proposed standard!

Nearly 2 years later, I have continued to work almost exclusively with SIP since that day in my position with WorldCom, giving seminars and teaching the protocol to others. This book grew out of those seminars and my work on various Internet-Drafts.

This revolutionary protocol was also the discovery of a radical standards body—the Internet Engineering Task Force (IETF). Later, I attended my first IETF meeting, which was for me a career changing event. To interact with this dedicated band of engineers and developers, who have quietly taken the Internet from obscurity into one of the most important technological developments of the late 20th century, for the first time was truly exciting.

Just a few short years later, SIP has taken the telecommunications industry by storm. The industry press contains announcement after announcement of SIP product and service support from established vendor startups, and from established carriers. As each new group and company joins the dialog, the protocol has been able to adapt and grow without becoming unwieldy or overly complex. In the future, I believe that SIP, along with a TCP/IP stack, will find its way into practically every intelligent electronic device that has a need to communicate with the outside world.

With my telecommunications background, it is not surprising that I rely on telephone examples and analogies throughout this book to explain and illustrate SIP. This is also consistent with the probability that telecommunications is the first widely deployed use of the protocol. SIP stacks will soon be in multimedia PCs, laptops, palmtops, and in dedicated SIP telephones. The protocol will be used by telephone switches, gateways, wireless devices, and mobile phones. One of the key features of SIP, however, is its flexibility; as a result, the protocol is likely to be used in a whole host of applications that have little or nothing to do with telephony. Quite possibly one of these applications, such as instant messaging, may become the next "killer application" of the Internet. However, the operation and concepts of the protocol are unchanged regardless of the application, and the telephone analogies and examples are, I feel, easy to follow and comprehend.

The book begins with a discussion of the Internet, the IETF, and the Internet Multimedia Protocol Stack, of which SIP is a part. From there, the protocol is introduced by examples. Next, the elements of a SIP network are discussed, and the details of the protocol in terms of message types, headers, and response codes are covered. In order to make up a complete telephony system, related protocols, including Session Description Protocol (SDP) and Real-Time Transport Protocol (RTP), are covered. SIP is then compared to another signaling protocol, H.323, with the key advantages of SIP highlighted. Finally, the future direction of the evolution of the protocol is examined.

Two of the recurring themes of this book are the simplicity and stateless nature of the protocol. Simplicity is a hallmark of SIP due to its text-encoded, highly readable messages, and its simple transactional model with few exceptions and special conditions. Statelessness relates to the ability of SIP servers to store minimal (or no) information about the state or existence of a media session in a network. The ability of a SIP network to use stateless servers that do not need to record transactions, keep logs, fill and empty buffers, etc., is, I believe, a seminal step in the evolution of communications systems. I hope that these two themes become apparent as you read this book and learn about this exciting new protocol.

The text is filled with examples and sample SIP messages. I had to invent a whole set of IP addresses, domain names, and URLs. Please note that they are all fictional—do not try to send anything to them.

I would first like to thank the group of current and former engineers at WorldCom who shared their knowledge of this protocol and gave me the opportunity to author my first Internet-Draft document. I particularly thank Henry Sinnreich, Steve Donovan, Dean Willis, and Matt Cannon. I also thank Robert Sparks, who I first met at the first seminar on SIP that I ever presented. Throughout the whole 3-hour session I kept wondering about the guy with the pony tail who seemed to know more than me about this brand new protocol! Robert

and I have spent countless hours discussing fine points of the protocol. In addition, I would like to thank him for his expert review of this manuscript prior to publication—it is a better book due to his thoroughness and attention to detail. I also thank everyone on the IETF SIP list who has assisted me with the protocol and added to my understanding of it.

A special thanks to my wife Lisa for the terrific cover artwork and the cool figures throughout the book.

Finally, I thank my editor Jon Workman, the series editor and reviewer, and the whole team at Artech for helping me in this, my first adventure in publishing.

Contents