

电子工程系列丛书 (影印版)

# Stream Control Transmission Protocol(SCTP) A Reference Guide

## 流控制传输协议 SCTP参考指南



RANDALL R. STEWART · QIAOBING XIE

电子工程系列丛书（影印版）

# **Stream Control Transmission Protocol (SCTP)**

**A Reference Guide**

## **流控制传输协议 SCTP 参考指南**

Randall R. Stewart  
Qiaobing Xie

清华大学出版社

EISBN 0-201-72186-4

Stream Control Transmission Protocol(SCTP) A Reference Guide

Randall R. Stewart, Qiaobing Xie

Copyright © 2002 by Addison-Wesley

Original English Language Edition Published by Addison-Wesley

All Rights Reserved.

For sale in Mainland China only.

本书影印版由培生教育出版集团授权清华大学出版社在中国境内独家出版、发行，香港、澳门特别行政区和台湾地区除外。

未经出版者书面许可，不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有培生教育出版集团激光防伪标签，无标签者不得销售。

北京市版权局著作权合同登记号：图字：01-2002-4532

书 名：流控制传输协议 SCTP 参考指南

作 者：Randall R. Stewart, Qiaobing Xie 著

出版者：清华大学出版社（北京清华大学学研大厦，邮编 100084）

<http://www.tup.tsinghua.edu.cn>

印刷者：北京牛山世兴印刷厂

发行者：新华书店总店北京发行所

开 本：787×960 1/16 印张：23.5

版 次：2003 年 1 月第 1 版 2003 年 1 月第 1 次印刷

书 号：ISBN 7-900643-95-8

印 数：0001~4000

定 价：49.00 元（含光盘）

# 出版前言

电子工程是信息科学的基础。高等学校新的教学要求指出：计算机专业和电子专业的学生应相互学习并渗透到彼此的专业领域，拓宽知识面，以适应信息技术飞速发展的时代。培养通晓相关专业领域知识的人才，已成为 21 世纪理工科教育的迫切要求。为此，我们挑选与信息科学、电子学有关的国外优秀著作，组成“电子工程系列丛书（影印版）”，奉献给国内读者。1999 年我们曾推出了奥本海姆的《信号与系统（第 2 版）》、奥法尼德斯的《信号处理导论》和拉贝的《数字集成电路》。这三本影印版图书获得了读者的广泛支持。本世纪我们将继续进行这项工作。根据读者的意见，今后我们出版的影印版图书，其开本尺寸不再缩小，基本保持其原版开本尺寸。

我们希望，这套丛书能为国内高校师生、工程技术人员以及科研单位的工作人员提供新的知识和营养，也衷心期待着读者对我们一如既往的支持。

清华大学出版社  
计算机教育出版事业部  
2002 年 10 月

To my parents, Jack J. and Betty M. Stewart. Their guidance, love, and direction throughout my life has shown me that you can do anything you set your mind to!

—Randall R. Stewart

To my parents, Weillie-Zeng and Shaojiang Xie.

—Qiaobing Xie

---

# Foreword

---

Over the years many people have felt that the widely available transport protocols in the TCP/IP protocol suite were not well matched for their particular applications. Many mechanisms have been proposed to change TCP—even going as far as suggesting a “TCP next generation”—to meet a wider set of requirements. However, the community has a well-founded reluctance to accept significant changes into the standard protocols for a number of reasons (for example, robustness issues and complexity issues). This makes the fact that the Stream Control Transmission Protocol (SCTP) has been developed and standardized even more impressive. SCTP is not simply a valuable technical achievement; it also shows the IETF’s ability to work with a particular community to solve problems in standard ways, resulting in a protocol that is general and likely applicable beyond the envisioned applications.

In this book Randall Stewart and Qiaobing Xie provide a valuable resource for those wishing to learn about and understand SCTP. This book presents SCTP on several levels, from the original motivations of the protocol, to the design decisions made, to a discussion of a working implementation. This book can be used for everything from gleaning a high-level overview of SCTP and the design process to gaining a nuts-and-bolts understanding of the interactions present in the protocol and their purpose.

This book can provide a wonderful companion to the SCTP specification for those implementing the protocol. While the specification (Stewart et al. 2000) outlines the exact syntax and semantics of SCTP, this book provides readers with a behind-the-scenes look at why certain features were included and the motivations behind the design decisions that were taken. The additional insights will likely help implementers better understand the protocol and therefore lead to better implementations of SCTP. In addition, the book contains an open-source version of SCTP that readers can reference. The concepts under discussion are not simply abstract; they are also shown in great detail.

Finally, this book strengthens the traditional way the Internet protocols have been documented. Traditionally the community writes a specification for a protocol, which is simply the final result of the design effort. As time passes, the community's memory of that process fades and some of the key issues discussed are lost. This book, on the other hand, offers inside insights (by two of the SCTP designers) into the design of the protocol and discusses why the protocol operates as it does. This insight takes much of the "black magic" out of SCTP and makes the protocol more accessible to those future protocol engineers who may be engaged in using and extending SCTP in the future but were not involved in the original SCTP standardization effort. More generally, the design process involved in writing SCTP and discussed in this book provides a valuable reference to those designing new protocols so they can understand the decision-making process and pitfalls that SCTP went through, and thus avoid them in their own work.

This book represents a significant step forward in the community's documentation of transport protocols and their design.

—Mark Allman  
Computer Scientist  
BBN Technologies  
August, 2001

---

# Preface

---

**T**his book describes, from the two primary designers' view, the design and operation of the Stream Control Transmission Protocol (SCTP)—a new general-purpose IP transport protocol recently standardized by the Internet Engineering Task Force (IETF).

For those who are familiar with TCP/IP networking, SCTP can be simply viewed as a “super-TCP,” that is, a TCP-like, general-purpose reliable data transport protocol with an enhanced set of capabilities.

SCTP is a key piece of the puzzle for making IP a truly viable technology choice in building the next-generation commercial grade infrastructure for telecommunications and e-commerce. For instance, SCTP has been required by the 3rd Generation Partnership Project (3GPP) to carry call signaling traffic in the third-generation cellular systems, and recommended by the IETF to carry the mission-critical Authentication, Authorization, and Accounting (AAA) messages in any future IP service networks.

*Telecommunications Magazine* selected SCTP as one of the “10 Hottest Technologies of Year 2001” in a cover story in its May 2001 issue.

## ***Why Do We Need SCTP?***

The basic design of TCP/IP has worked remarkably well and remained largely unchanged. All the while the Internet has gone through its exponential growth, from a network of only a few nodes connecting a few university campuses, to a truly global Internet with tens of millions of hosts.

The Internet's explosive growth, witnessed by the nineties, continues into the new millennium. At the same time, IP technology has been steadily transforming itself into a cornerstone of the new information-era economy. One of the most visible results of this transformation is that IP technology is beginning to see more and more commercial uses. This is in sharp contrast to the seventies and eighties, when almost the only mention of TCP/IP was heard in academic and research communities.



It is inevitable that this commercialization process has brought with it new requirements to the technology. The original design of TCP, which was conceived and incubated in a mostly academic and research environment, has started to show its limitations in meeting some of these new commercial requirements.

It is this commercialization challenge on IP technology that has brought on the birth of SCTP.

When we first brought the original idea of SCTP (then called MDTP) to the IETF, we were trying to create a protocol to solve a very particular and practical problem; that is, how to transport telephony signaling messages using IP technology in commercial-grade systems. But, at the end of the process, the success of SCTP has grown far beyond our widest expectations.

### ***Readers***

This book is written with two different groups of readers in mind.

The first group consists of those who want to understand SCTP. These readers may include programmers who need to work with SCTP, managers who want to make informed decisions about SCTP, students who want to learn the basics of SCTP, etcetera. For them, this book can serve as an easy-to-read alternative to the IETF standards specification.

In this book they will find not only detailed explanations of all the basic parts and operations of the protocol, often accompanied by a good deal of examples and diagrams, but also a thorough item-by-item comparison between functions of SCTP and TCP/UDP. Some of the chapters also give this group of readers guidance on which parts of the chapter they can safely skip without hindering their understanding of the general protocol operation.

The second group of readers we have in mind consists of the implementers of SCTP. For them, this book can be used as an annotation to the IETF standards specification. It is our hope that they will find some helpful clarifications of certain complex internal mechanisms of the protocol through the examples and illustrations we put in the book. Moreover, the recounts we have in the book on many important IETF debates and consensus-building processes may give them some insight into the protocol design. This in turn may improve their understanding over the rationale behind those same key design decisions.

### ***Prerequisites***

We assume that the readers have preliminary knowledge about TCP/IP and IP networking. (An excellent book for this is W. Richard Stevens' and Gary R. Wright's *The TCP/IP Illustrated, Volume 1: The Protocols*, published by Addison-Wesley.) Moreover, some experience with C programming and Unix systems will be very helpful, though it is not an absolute requirement.

For those who would like to refresh their knowledge of IP networking, we have a short review of the IP networking basics at the beginning of Chapter 1.

### **Coverage**

This book covers all aspects of SCTP standards as defined in the IETF's *Request for Comments 2960 (RFC2960)* (Stewart et al. 2000). We also include a preview of the proposed SCTP sockets API (application programming interface) extension, which is a work in progress in the IETF.

A CD-ROM containing the open-source user-space SCTP reference implementation is included in this book. Readers who want to read the reference implementation code can find a chapter at the end of the book explaining the overall design and internal structures of the code. This code has been compiled and tested with FreeBSD 4.2, Linux 2.2.17 and 2.4.1, and Solaris 2.8.1 operating systems. Updates and new releases of this open-source SCTP reference implementation can be found at <http://www.sctp.org>.

For those interested in doing experiments with the SCTP reference implementation, a good tool to use to capture and analyze SCTP traffic is the Ethereal Network Analyzer, which is freely downloadable from <http://www.ethereal.com>.

### **Acknowledgments**

Never in our estimation did we expect to participate in writing an IETF *Request for Comments (RFC)* and subsequently write a book on SCTP, a major transport protocol of the future Internet.

It has been a privilege and an honor to work with the many fine individuals in the IETF SIGTRAN working group. It is through their comments, thoughts, and opinions that SCTP was crafted and refined. Without all of their input, SCTP would be less. Our sincere gratitude therefore goes out to all of our coauthors of *RFC2960* and to many other individuals who contributed greatly in commenting on and reviewing the design.

A special thanks to the SIGTRAN working group chair, Lyndon Ong, and the IETF Transport Area directors, Scott Bradner and Vern Paxson. They have provided insightful guidance and encouragement to “think of the bigger picture” and look into the future. Their superb management and masterful consensus-building skill throughout the process helped speed the specification through the IETF process.

This is the first time for both of us writing a technical book. Undoubtedly our inexperience does make this book-writing adventure a more challenging one. We especially thank our editing staff, Emily Frey, Karen Gettman, Mary Hart, Robert Kern, Jeannine Kolbush, and Elizabeth Ryan, for their step-by-step guidance and patience.

We also thank greatly our technical reviewers, Lode Coene, Muckesh Kacker, John Loughney, Ken Morneault, Kacheong Poon, Ian Rytina, Hanns Juergen Schwarzbauer, Chip Sharp, and Lars Viklund, for providing their different views on various issues, pointing out our numerous editorial oversights and inconsistencies, and keeping us honest by catching technical mistakes.

R.R.S.: A special note of thanks also to my wife Sandra for putting up with me, my late hours, and my constant “calling in to get e-mail.” Dear, you are truly my soul mate, and without you I would be lost.

Q.X.: I would especially like to thank my wife Qing for her understanding and patience throughout the writing of this book.

—Randall R. Stewart  
Crystal Lake, Illinois

—Qiaobing Xie  
Palatine, Illinois

---

# Contents

---

<i>FOREWORD</i>	<i>XV</i>
<i>PREFACE</i>	<i>XVII</i>
<b>1 Introduction and Overview</b>	<b>1</b>
1.1 IP Networking Basics	1
1.1.1 How Messages Are Delivered in an IP Network	2
1.1.2 IP Protocols	5
1.1.3 IP Multi-homing	6
1.2 What Is SCTP?	6
1.2.1 Where Does SCTP Fit in the IP Architecture?	7
1.3 Motivation for Developing SCTP	8
1.3.1 TCP Limitations	8
1.3.2 UDP Limitations	10
1.3.3 SCTP Enhancements over TCP and UDP	10
1.4 A Short History of SCTP Development	11
1.4.1 Early Works Before the IETF and MDTP	11
1.4.2 IETF Refinements	13
1.5 Major General SCTP Issues Debated in the IETF	14
1.5.1 Do We Really Need a New Transport Protocol?	14
1.5.2 Over UDP Versus Over IP	16
1.6 Organization of this Book	17
1.7 Summary	19
1.8 Questions	19
<b>2 SCTP Endpoint and Association</b>	<b>21</b>
2.1 SCTP Transport Address	21
2.2 SCTP Endpoint	22
2.2.1 Multi-homed SCTP Endpoint	22

2.3	SCTP Association	24
2.4	Operation of an SCTP Association	26
2.4.1	Functional View of an Association	26
2.4.2	Functional Responsibilities of an SCTP Endpoint	28
2.4.3	Operation States of an Association	30
2.5	IETF Debate and Issues	33
2.6	Summary	33
2.7	Questions	34
<b>3</b>	<b>Format of SCTP Messages</b>	<b>35</b>
3.1	Basic Layout—The Bits on the Wire	35
3.1.1	The SCTP Common Header	36
3.1.2	Elements in a Chunk	37
3.1.3	Defined Chunk Types and Extensibility	38
3.1.4	General Rules When Dealing with Chunks	41
3.1.5	Chunk Parameters	42
3.1.6	Other Considerations	44
3.2	Basic Chunk Descriptions	44
3.2.1	The INIT Chunk	44
3.2.2	The INIT-ACK Chunk	50
3.2.3	The COOKIE-ECHO Chunk	54
3.2.4	The COOKIE-ACK Chunk	55
3.2.5	The DATA Chunk	55
3.2.6	The SACK Chunk	58
3.2.7	The HEARTBEAT Chunk	60
3.2.8	The HEARTBEAT-ACK Chunk	61
3.2.9	The SHUTDOWN chunk	62
3.2.10	The SHUTDOWN-ACK Chunk	63
3.2.11	The SHUTDOWN-COMPLETE Chunk	64
3.2.12	The ERROR Chunk	65
3.2.13	The Error Cause Parameters	65
3.2.14	The ABORT Chunk	72
3.3	IETF Debate and Issue	73
3.4	Summary	75
3.5	Questions	75

<b>4</b>	<b>Setup of an Association</b>	<b>77</b>
4.1	The INIT Chunk	79
4.1.1	Formulating the Chunk	79
4.1.2	INIT Parameters	80
4.1.3	Timer and Retransmission for INIT	83
4.2	The INIT-ACK Response	83
4.2.1	Formulating the Response	84
4.2.2	INIT-ACK Parameters	85
4.2.3	Preventing Resource Attacks	89
4.3	The COOKIE-ECHO Chunk	89
4.3.1	Formulating the Chunk	91
4.3.2	Bundling Data with the COOKIE-ECHO	91
4.3.3	Timer and Retransmission for COOKIE-ECHO	91
4.4	The COOKIE-ACK Chunk	92
4.4.1	Validating and Unpacking the Cookie	93
4.4.2	Formulating the COOKIE-ACK	95
4.4.3	Bundling DATA and SACK with COOKIE-ACK	96
4.5	Address Usage When Starting an Association	97
4.6	Chunk Rejections During an Association Setup	101
4.6.1	When Is an INIT Not Acceptable?	102
4.6.2	When Is a Cookie Not Worth Eating?	102
4.7	Handling Unexpected Initialization Chunks	103
4.7.1	Association Tie-Tags and Their Values	104
4.7.2	Handling Overlapping Initialization Chunks	105
4.7.3	Case Study 1: Initialization Collision	110
4.7.4	Case Study 2: Peer Restart	119
4.7.5	A Lost COOKIE-ACK	121
4.8	IETF Debate and Issues	122
4.9	Summary	124
4.10	Questions	124
<b>5</b>	<b>User Data Transfer</b>	<b>125</b>
5.1	User Message Transfer Overview	125
5.1.1	How User Messages Are Sent	125
5.1.2	How User Messages Are Received	129
5.1.3	Summary of Data Transfer Overview	131
5.2	Obedying the Transmission Rules	132
5.2.1	Congestion Window and Receiver Window	132

	5.2.2 Rules for Data Transmission	132
5.3	Acknowledgment Rules	137
	5.3.1 Where Should a SACK Be Sent?	137
	5.3.2 When Should a SACK Be Sent?	138
	5.3.3 What Should Be Reported in a SACK?	142
	5.3.4 Revoking an Acknowledgment by the Receiver	146
	5.3.5 Processing of Received SACK by Data Sender	148
5.4	Management of Retransmission Timer	151
	5.4.1 Determine Retransmission Timeout (RTO)	151
	5.4.2 Retransmission Timer Rules	154
	5.4.3 Handling Retransmission Timer Expiration	156
5.5	Multi-homed Data Sender and Receiver	158
	5.5.1 Selection of Primary Transport Address	158
	5.5.2 Transferring Data between Multi-homed Endpoints	159
	5.5.3 Sending a SACK on Duplicate Data	161
5.6	Stream Identifier and Stream Sequence Number	162
	5.6.1 Ordered and Unordered Message Delivery	164
5.7	Passing a Payload Protocol Identifier	165
5.8	Fragmentation and Bundling of User Messages	166
	5.8.1 When Should User Messages Be Fragmented?	166
	5.8.2 How Should a User Message Be Fragmented?	167
	5.8.3 Reassembling a Fragmented Message at the Data Receiver	168
	5.8.4 User Message Bundling	169
5.9	Adler-32 Checksum	171
	5.9.1 Generation of the Adler-32 Checksum by the Data Sender	171
	5.9.2 Validation of the Adler-32 Checksum by the Data Receiver	172
5.10	Error Handling	172
	5.10.1 Handling a Data Chunk with an Invalid Stream Identifier	172
	5.10.2 Handling an Empty Data Chunk	173
	5.10.3 Handling Out-of-State Data	173
5.11	IETF Debate and Issues	173
5.12	Summary	175
5.13	Questions	175

<b>6</b>	<b>Congestion Control and Avoidance</b>	<b>177</b>
6.1	The Cause of Network Congestion Collapse	177
6.2	Basic Concept of Congestion Control	178
6.3	SCTP Congestion Control Algorithm	180
6.3.1	How Does Your cwnd Grow?	183
6.3.2	How Does Your cwnd Shrink?	185
6.4	Interesting Details of the Algorithms	186
6.5	Path MTU Discovery	193
6.5.1	A Refresher on How PMTU Works	193
6.5.2	PMTU Discovery in SCTP	195
6.6	Explicit Congestion Notification	196
6.7	IETF Debate and Issues	197
6.8	Summary	199
6.9	Questions	200
<b>7</b>	<b>Failure Detection and Recovery</b>	<b>201</b>
7.1	The Types of Failure SCTP Detects	201
7.1.1	Unreachable Destination Address	201
7.1.2	Unreachable Peer Endpoint	202
7.2	How to Detect an Unreachable Address	203
7.2.1	Using HEARTBEAT Chunks to Monitor Reachability	204
7.2.2	Address-Unreachability Detection Algorithm	205
7.2.3	More on Using the SCTP HEARTBEAT	209
7.2.4	Handling an Unreachable Primary Destination Address	213
7.3	How to Detect an Unreachable Peer	213
7.3.1	When an Association Enters the Dormant State	216
7.4	Fault Resilience Communication and Routing Configuration	218
7.4.1	Maximizing Path Diversity	218
7.4.2	Asymmetric Multi-homing Configuration	220
7.4.3	Effects of Source Address Selection	222
7.5	IETF Debate and Issues	225
7.6	Summary	226
7.7	Questions	226



<b>8</b>	<b>Auxiliary SCTP Packet Handling Functions</b>	<b>229</b>
8.1	Handling Out-of-the-Blue SCTP Packets	229
8.2	SCTP Packet Verification Tag Rules	231
8.2.1	Basic Verification Tag Rules	232
8.2.2	Special Verification Tag Rules	232
8.3	IETF Debate and Discussion	235
8.4	Summary	236
8.5	Questions	236
<b>9</b>	<b>Close of an Association</b>	<b>237</b>
9.1	The Graceful Shutdown	238
9.1.1	Sending SHUTDOWN	239
9.1.2	Receiving SHUTDOWN	241
9.1.3	Sending SHUTDOWN-COMplete	242
9.1.4	Race Conditions and Lost Messages	244
9.2	The Abortive Shutdown	248
9.2.1	ABORT Usage	248
9.2.2	Sending an ABORT	248
9.2.3	Receiving an ABORT	249
9.3	IETF Debate and Issues	250
9.4	Summary	250
9.5	Questions	251
<b>10</b>	<b>IANA and Future Extension of SCTP</b>	<b>253</b>
10.1	SCTP Port Number Assignment	253
10.2	IETF Chunk Extensions	254
10.3	IETF Parameter Extensions	255
10.4	IETF Extensions to Error Causes	255
10.5	Payload Protocol Identifier	256
10.6	Summary	256
<b>11</b>	<b>A Sockets API for SCTP</b>	<b>257</b>
11.1	A Quick Review of the TCP Sockets API	258
11.1.1	The Server-Side Application	258
11.1.2	The Client-Side Application	259
11.2	The TCP-Style SCTP Sockets API Extension	261
11.3	The UDP-Style SCTP Sockets API Extension	262