SCTP: An innovative transport layer protocol for the web

(Position paper)

P. Natarajan, J. Iyengar, P. Amer, & R. Stewart
HTTP over TCP

• Transmission Control Protocol (TCP) has been the default transport for HTTP.
• HTTP/TCP Concerns
  – Head-of-line (HOL) blocking
  – Vulnerability to network failures
  – Vulnerability to SYN DoS attacks
HOL blocking in TCP

HOL BLOCK!

objects in send buffer

6 5 4 3 2 1

delivered to application

6 5 4 3 2 1

receive buffer

6 5 4 3 2 1

retransmission

TCP Connection

loss

Web server

P.E.L.

Web client

Cisco Systems
SCTP multistreaming avoid HOL blocking

objects in send buffer

delivered to application

receive buffer

retransmission

Web server

Web client

SCTP Association

loss

Stream 1
TCP work-around to mitigate HOL blocking

• How?
  – Multiple persistent TCP connections to transfer independent web objects

• Problems
  – Possible HOL blocking within one TCP connection
  – No shared sequence space => Less robust to loss detection and recovery
  – Increased load on web server
  – Increased connection establishment latency during SYN losses.
  – Aggressive behavior during congestion
TCP: Network fault-(In)tolerance
SCTP: Transport layer multihoming

SCTP Association: \((\{A_1, A_2\}, \{B_1, B_2\})\)

SCTP Failure Detection & Failover
TCP SYN Flooding Attack

Internet

Spoofed SYNs

128.3.4.5
SYN 130.2.4.15

192.10.2.8
SYN 228.3.14.5

221.3.5.10
SYN 190.13.4.1

TCP web server

Flooded!!
SCTP Association setup avoids SYN flooding attack

Spoofed INITs

Internet

128.3.4.5 INIT 130.2.4.15

192.10.2.8 INIT 228.3.14.5

221.3.5.10 INIT 190.13.4.1

SCTP web server

Process INIT
SCTP: Four-way Association setup

INIT

INIT-ACK (StateCookie)

COOKIE-ECHO (StateCookie); DATA

COOKIE-ACK; SACKs

DATA

NO TCB

TCB
HTTP/SCTP streams: Design

HTTP Client

"HTTP REQUEST"
"HTTP RESPONSE"

(to Read m)

SCTP

Stream ID: m
"HTTP REQ"
Stream ID: m
"HTTP RESP"

HTTP Server

"HTTP REQUEST"
"HTTP RESPONSE"

Write (to stream m)

SCTP Association

Stream ID: m
"HTTP REQ"
Stream ID: m
"HTTP RESP"

SCTP

Stream ID: m
"HTTP RESP"
Stream ID: m
"HTTP RESP"
HTTP/SCTP Implementation

• Apache 2.0.55
• Firefox 1.6a
It Works!

HTTP over TCP

HTTP over SCTP (multistreaming)

TCP TPDU

SCTP TPDU

Object 1

Object 2

Object 3

Object 4

Object 5

time

= PDU lost
Other SCTP features

• Preservation of Message Boundaries
• Partial Reliability Extension (PR-SCTP)
  – Timed reliability: Attempt for reliable transmission only within a time period.
  – Example: Online game client use PR-SCTP to transmit player’s coordinates. Old coordinates discarded when newer ones available.
• Unordered data delivery
  – 1 SCTP association to transmit both ordered and unordered data
  – Vs. UDP: Unordered data transmitted reliably.
• SCTP shim layer
  – Between application and transport layer.
  – No code change to app. Transparently converts app’s TCP calls to corresponding SCTP calls.
Current status

- **Home**: IETF TSVWG (Transport Services Working Group)
  - IETF recognizes broader scope
  - Proposed Standard - RFC2960

<table>
<thead>
<tr>
<th>Interops (8)</th>
<th>Date</th>
<th>#Impl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Munich</td>
<td>6/00</td>
<td>12</td>
</tr>
<tr>
<td>Research Triangle Park</td>
<td>10/00</td>
<td>22</td>
</tr>
<tr>
<td>Sophia Antipolis</td>
<td>4/01</td>
<td>19</td>
</tr>
<tr>
<td>San Jose (Connectathon)</td>
<td>2/02</td>
<td>6</td>
</tr>
<tr>
<td>U of Essen (Germany)</td>
<td>9/02</td>
<td>20</td>
</tr>
<tr>
<td>U of Delaware</td>
<td>6/03</td>
<td>11</td>
</tr>
<tr>
<td>U of Muenster (Germany)</td>
<td>7/04</td>
<td>14</td>
</tr>
<tr>
<td>Vancouver</td>
<td>8/06</td>
<td>??</td>
</tr>
</tbody>
</table>

- **Supported by industry**:
  - **Implementations**: AIX, FreeBSD, NetBSD, DragonFly BSD, Linux, QNX, Solaris, True64, IOS (Cisco Routers), Mac OS, Windows (user space), more…
References - RFCs

- RFC 2960 – Stream Control Transmission Protocol
- RFC 3257 – SCTP Applicability Statement
- RFC 3286 – An introduction to SCTP
- RFC 3309 – SCTP Checksum Change
- RFC 3436 – Transport Layer Security over SCTP
- RFC 3554 – On the Use of SCTP with IPsec
- RFC 3758 – SCTP Partial Reliability Extension
- RFC 4460 – SCTP Specification Errata and Issues
References – Internet Drafts

• SCTP (BIS)
  – draft-ietf-tsvwg-2960bis-01.txt
• Sockets API Extensions for SCTP
  – draft-ietf-tsvwg-sctpsocket-12.txt
• SCTP Dynamic Address Reconfiguration (Add-IP)
  – draft-ietf-tsvwg-addip-sctp-14.txt
• SCTP Packet Drop Reporting (Pkt-Drop)
  – draft-stewart-sctp-pktdrprep-04.txt
• Authenticated Chunks for SCTP (Auth)
  – draft-tuexen-sctp-auth-chunk-02.txt
References - Books


  – chapter 2: The Transport Layer: TCP, UDP, and SCTP
  – chapter 9: Elementary SCTP Sockets
  – chapter 10: SCTP Client/Server Example
  – chapter 23: Advanced SCTP Sockets

  – chapter 13: SCTP
References - Papers

• Caro Jr. et al, “**SCTP: A Proposed Standard for Robust Internet Data Transport**”, IEEE Computer 36(11), 11/03
• Stewart & Amer, **Internet Society Brief 17**
• Univ of Delaware Protocol Engineering Lab (**PEL**)
References – Online

• http://www.sctp.org
  – Also reachable with HTTP over SCTP!
• http://www.ietf.org/html.charters/tsvwg-charter.html
  – All current work on SCTP is done in the IETF TSVWG
• sctp-impl on mailer.cisco.com
  – Note for Cisco audience: this is an external list
Questions