Reliability and Availability in Stream Control Transport Protocol (SCTP)

Research Seminar on
Real Time and High Availability
Autumn 2001

by Laila Daniel
on 21st Nov. 2001

Stream Control Transmission Protocol (SCTP)

- SCTP is a reliable transport protocol operating on top of a connectionless packet network such as IP
- Exists at an equivalent level as TCP and UDP
- SCTP is designed by IETF SIGTRAN Working Group
- SCTP is an IETF Proposed Standard
- SCTP is specified in RFC 2960 - released in Oct. 2000
**SCTP Offers**

- Acknowledged, error free, non-duplicate transfer of user data
- Data fragmentation to conform to discovered path MTU size
- Sequenced delivery of user messages within multiple streams with an option for order-of-arrival delivery of individual user messages
- Optional bundling of user messages into a single SCTP packet
- Network level fault tolerance through multihoming

**SCTP vs TCP**

- Similarities to TCP
  - Connection Oriented
  - Reliable
  - Uses IP (or other unreliable packet network service)
  - Flow Control and Congestion Control measures are similar (TCP friendly)
SCTP vs TCP

- Differences to TCP
  - Out of sequence delivery of data possible
  - Multiple streams of chunks - not one byte stream
  - Stream sublayer to avoid head-of-line blocking
  - Selective ACK (SACK) is part of the standard
  - Multihoming for enhanced network fault tolerance
  - Enhanced security

SCTP Application Examples (1/3)

- SCTP was originally designed to transport PSTN (Public Switched Telephone Network) signalling messages over IP networks.
- SCTP can be applied to transport messages of classical telecommunications signalling protocols like:
  - SS7 (Signalling System No. 7)
  - DSS 1 (Digital Subscriber System No. 1 - ISDN)
  - DSS 2 (Digital Subscriber System No. 2 - B-ISDN)

over IP
SCTP Application Examples (2/3)

- Interconnection of PSTN islands via IP networks (Virtual Trunking)
- Access from PSTN to IP based Service Control Points (SCP) for Intelligent Network Services
- Access to IP based location registers for mobility management in mobile networks
- Signalling transport in the UMTS (Universal Mobile Telecommunications Systems) Radio Access Networks

SCTP Application Examples (3/3)

- Voice over IP
- Can be used directly in any IP-based Network
- Alternative to TCP and UDP in some applications
Performance requirements of SS7 signalling scheme

- Error rate < 1 in $10^{10}$ message signal units
- Message loss rate < 1 in $10^7$ messages
- The availability of any signalling relation (a path between two communicating points) has to be at least 0.99998 corresponding to a downtime of at most 10 min/year
- Out of sequence / Duplication of message < 1 in $10^{10}$
- Message transfer time < 100 msec

TCP performance problems

- TCP provides reliable data transfer and strict order-of-transmission delivery of data. Some applications need reliable data transfer without sequenced delivery
- Byte stream oriented nature requires applications to add their own record marking
- Using TCP sockets it is difficult to provide highly available data transfer using multihomed hosts
- TCP is vulnerable to denial of service attacks such as SYN attacks
Functional View of the SCTP Transport Service

An SCTP Association
SCTP Packet Format

SCTP Association Setup

State of client | Exchanged Chunk | State of server
--- | --- | ---
Closed | Tag = 0; INIT[Tag = Tag_C] | Closed
Cookie-Wait | Tag = Tag_C; INIT-ACK[Tag = Tag_S; cookie] | Closed
Cookie-Wait | Tag = Tag_S; COOKIE-ECHO [cookie] | Established
Cookie-Echoed | Tag = Tag_C; COOKIE-ACK | Established
Established | | Established
Real Time support in SCTP

- Multistreaming
- Multihoming
- Facility for unordered data
- Sending one SACK per received packet once gap is observed
- Application can specify the life time of time sensitive signalling messages

Format of a Data Chunk

<table>
<thead>
<tr>
<th>Type = 0</th>
<th>ReservedUBE</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transmission Sequence Number - TSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stream Identifier S</td>
</tr>
<tr>
<td>Payload Protocol Identifier</td>
</tr>
</tbody>
</table>

User Data (seq n of Stream S)
SCTP’s Reliable Data Transfer with Flexible Data Delivery

- SCTP creates independence between data transmission and data delivery
- Each Data Chunk uses two sets of Sequence Nos.
  - TSN (Transmission Seq. No) for
    - the transmission of messages
    - detection of message loss
  - Stream ID / Stream Seq. No. for
    - determining the sequence of delivery of received data

SCTP operates on two levels
- Without reconstructing the initial sequence, every correctly received data chunk is delivered to a second level
- The second level provides a flexible delivery mechanism based on several independent streams of packets within an association
Reliable Data Transfer

- Every data chunk has a Transport Sequence Number (TSN). ACKs returned to the sender are based on TSNs.
- 32 bit Checksum (Fletcher-16) is used to validate the packets.
- The receiver acknowledges each received packet by sending a SACK chunk reporting all gaps.
- Whenever the sender gets four consecutive gap reports for the same data chunk, this data chunk is immediately retransmitted (Fast Retransmit).

Flexible Data Delivery

- Flexible data delivery is achieved using Multi-streaming. Stream is a uni-directional logical channel established from one to another associated SCTP endpoint, within which all user messages are delivered in sequence except for those submitted to the unordered delivery service.
- At the association setup the number of available streams per association can be exchanged between peer entities.
Flexible Data Delivery (contd...)

- Multi-streaming allows data to be partitioned into multiple streams that have the property of being delivered independently
- SCTP assigns a Stream Sequence Number (SSN) within each stream to each message passed to it by the server.
- At the receiving side SCTP assures that messages are delivered to the SCTP user in-sequence within a given stream (Partial in-sequence delivery)
- While one stream may be blocked from delivery for the next in-sequence user message, delivery from other streams may proceed
- Avoids head-of-line blocking between independent streams of packets within an association

Flexible Data Delivery (contd...)

- SCTP allows to mark packets for out-of-order delivery (i.e., order-of-arrival)
- Useful when important messages which have to bypass others.
Security Mechanisms in SCTP

- Verification Tag
- Cookie
- Use of IPSec

Fault Management
Monitoring of path states

- Multihoming allows to select from several transmission paths
- Heartbeat messages
- Retransmissions
SCTP Multi-Homing

- Multi-homing is the ability for a single SCTP endpoint to support multiple IP addresses.
- With multihoming traffic from one node to another travels on physically different paths if different destination IP address are used, associations become tolerant against physical network failures.
- The endpoints can exchange lists of addresses at the time of initiation of an association.

Heartbeat Messages

- Heartbeat chunks are sent to all idle destinations (i.e., all alternate addresses)
- ACK for a Heartbeat should be generated instantly by the associated end point
- A counter is maintained on the number of Heartbeats sent to an inactive destination without receipt of a corresponding Heartbeat ACK. If the counter exceeds a configured maximum, the address is declared to be inactive
- Heartbeats continue to be sent to inactive addresses until an ACK is got and the address is made active.
- Rate of sending of heartbeats is tied to RTO and the application
Retransmissions

- Retransmission occurs at
  - timeout of the retransmission timer
  - receipt of 4 SACKs for a data chunk
- A counter is maintained for number of retransmissions to a particular destination address without successful acknowledgement.
- When the counter exceeds a configured maximum the address is declared inactive
- A notification is given to the application and SCTP begins to use an alternate address for sending data

Performance Evaluation of SCTP

- Andreas Jungmair, University of Essen, Michael Schopp, Siemens AG, and Michael Tuxen, Siemens AG
  - ATM 2000 Conference Presentation June 2000, Heidelberg
Results – Benefits of out-of-sequence delivery

- Average delivery times vs. loss rate
- Assumed link delay
  - 50 ms
- In-sequence delivery
  - Delivery time grows due to resequencing delay
- Out-of-sequence delivery
  - Feature of SCTP that TCP does not have
  - No resequencing delay

Testbed for protocol evaluation

- Results were obtained in a testbed with Linux hosts
- SCTP reference implementation
  - this is the reference
  - others have to show similar behavior
- NISTNET War Emulator
  - packet delay,
  - loss,
  - bandwidth limitation
Results — SCTP throughput

- Plot shows protocol throughput as function of link delay
  - bandwidth/delay function
  - as seen by user level
- SCTP with varying chunk sizes (overhead)
- TCP
- both with 32 kBytes receiver window
- both assume their theoretical limits
- maximum link bandwidth is 100 kBytes/s

Results — Fair bandwidth sharing with TCP

- Diagram shows protocol throughput as function of time
- t=30 sec:
  - TCP instance starts bulk transfer
- t=130 sec:
  - TCP instance stops sending data
- Link is shared equally between SCTP instance and TCP instance!
SCTP API

- RFC 2960 includes a model of the primitives exchanged between the application and the SCTP layer
- Provides a socket-like API to simplify any migration of TCP or UDP applications to the use of SCTP

Implementation

- The SCTP implementation supports the current version of the protocol.
- The software is developed by
  - Siemens &
  - The Computer networking technology group of the University of Essen.
- It runs under
  - Linux, FreeBSD, Solaris and Mac OS
- It supports
  - IPv4 and IPv6
- It is published under the GNU public license.
Getting the source

• The SCTP source can be downloaded from http://www.sctp.de/download/index.html

• Support via mailing lists announce@sctp.de

Conclusion

• SCTP is a new reliable transport protocol
• Some superior features to TCP
• Open source implementation available
• SCTP can be deployed in existing IP networks
  - Fair bandwidth sharing among SCTP associations
  - Fair bandwidth sharing between TCP connections and SCTP associations
• Consequence for network design
  - Guaranteed bandwidth per association = available bandwidth / number of associations
RFCs related to SCTP

- RFC 2960 – Stream Control Transmission Protocol
- RFC 2719 - Architectural Framework for Signalling Transport
- RFC 3057 - ISDN Q.921-User Adaptation Layer.

Papers and Drafts related to SCTP

- L. Ong, J. Yoakum: An O verview of the SCTP, Internet Draft : draft-ong-sigtran-sctpover-01.txt
SCoTP related links (1/3)

- SCTP Primer  
  http://tdrwww.exp-math.uni-essen.de/pages/forschung/sctp_fb/
- SCTP for beginners  
  http://tdrwww.exp-math.uni-essen.de/pages/forschung/sctp_fb/sctp_links.html
- SCTP @ University of Delaware  
  http://www.cis.udel.edu/~iyengar/research/SCTP/
- SCTP @ Temple University  
  http://netlab.cis.temple.edu/SCTP/

SCoTP related links (2/3)

- Ethereal - http://www.ethereal.com  
  A graphical packet analyzer with SCTP support running under Linux, FreeBSD, Solaris, ... and Windows. IUA and
  The charter of the Signaling Transport (SIG TRAN) working group (WG) of the IETF. There are also links to all internet draft accepted as working groups documents.
SCTP related links (3/3)

- http://www.sctp.org
- http://www.sctp.de
- http://sourceforge.net/projects/lksctp/ - (Linux implementation)
- http://playground.sun.com/sctp - (for Solaris 2.8 implementation)