SCTP Streams

While TCP couples the reliable transfer of user data and the strict order-of-transmission delivery of such data, SCTP separates the reliable transfer of datagrams from the delivery mechanism. This makes it possible to adapt protocol usage to the specific needs of the applications using SCTP. Some applications may only need partial ordering of datagrams while others might even be satisfied with a reliable transfer that does not guarantee any in-sequence maintenance at all.

General Concepts

SCTP distinguishes different *streams* of messages within one SCTP association. This enables a delivery scheme where only the sequence of messages needs to be maintained per stream (partial in-sequence delivery) which reduces unnecessary head-of-line blocking between independent streams of messages. Furthermore, SCTP provides a mechanism for bypassing the sequenced delivery service, so that messages are delivered to the user of SCTP as soon as they are completely received (order-of-arrival delivery).

Flow control and congestion control in SCTP have been designed in a way which assures that SCTP traffic behaves in the network in the same way as TCP traffic does. This enables a seamless introduction of SCTP services into existing IP networks (see also Performance Evaluation of the Stream Control Transmission Protocol).

SCTP operates on two levels:

- Within an association the reliable transfer of datagrams is assured by using a checksum, a sequence number and a selective retransmission mechanism. Without taking the initial sequence into account, every correctly received data chunk is delivered to a second, independent level.

- The second level realises a flexible delivery mechanism which is based on the notion of several independent *streams* of datagrams within an association.

Detection of loss and duplication of data chunks is enabled by numbering all data chunks in the sender with the so-called Transport Sequence Number (TSN). The acknowledgements sent from the receiver to the sender are based on these sequence numbers.
Retransmissions are timer-controlled. The timer duration is derived from continous measurements of the round trip delay. Whenever such a retransmission timer expires, (and congestion control allows transmissions) all non-acknowledged data chunks are retransmitted and the timer is started again doubling its initial duration (like in TCP).
When the receiver detects one or more gaps in the sequence of data chunks, each received SCTP packet is acknowleged by sending a Selective Acknowledgement (SACK) which reports all gaps. The SACK is contained in a specific control chunk. Whenever the sender receives four consecutive SACKs on the same data chunk this data chunk is immediately retransmitted (fast retransmit). Most up-to-date operating systems already support a similar optional extension to TCP (see RFC 2018).

Flexible Datagram Delivery

The user of SCTP may assign each datagram to one of several *streams* within an association. When an association is set-up, the number of available streams per direction is exchanged between the peer entities. Within each stream, SCTP assigns independent Stream Sequence Numbers (SSN) to the user datagrams. These numbers are used at the receiver to determine the sequence of delivery. SCTP performs in-sequence delivery per stream (for all datagrams which are not marked for out-of-order delivery). This mechanism avoids head-of-line blocking between independent streams of datagrams within one association. With TCP, this could only be achieved by setting-up several connections (one per stream) which would lead to additional cost and overhead.

As already mentioned, SCTP allows to mark datagrams for order-of-arrival delivery. This could be used for important messages which may by-pass others, like e.g. transaction abort messages of an application. If no sequence maintenance is required, all datagrams could be marked accordingly.