

# Experimental studies of SCTP multi-homing

Thomas Ravier, Rob Brennan and Thomas Curran  
Teltec DCU, Dublin 9, Ireland

## Abstract

The Stream Control Transmission Protocol (SCTP) is a new reliable transport protocol operating on top of the Internet Protocol (IP). It is designed to transport telecommunications signalling traffic over an IP network and has been recently standardised by the Internet Engineering Task Force (IETF). “Multi-homing” is one of the key features of the protocol. It provides SCTP with a certain network level fault tolerance by the use of network address redundancy. This paper investigates the performance characteristics of multi-homed SCTP hosts through experimental studies. We briefly describe the role of SCTP and give an overview of its multi-homing feature. With the help of plots and network statistics gathered by instrumentation functions we added to the Reference Implementation of SCTP, we show SCTP performance and behaviour like the slow start phase. Through experimentation we observed the link changeover policy of the Reference Implementation of SCTP and the ability of SCTP to handle transmission of packets over a lossy link.

## 1. Introduction

Modern telecommunication networks depend heavily on the fast and reliable exchange of control information. Signalling (exchanging control information) between different network entities not only supports basic telecommunications services and features, it also enables the provisioning of advanced network services such as Intelligent Network (IN) services or mobility management in mobile communication networks. Quality of service criteria (e.g. the post-dialling delay in fixed and mobile networks) are also strongly influenced by the performance of the signalling system.

For many years, the signalling system number 7 (SS.7) [1] has been the dominant bearer of signalling traffic for telecommunications networks. The SS7 signalling network is logically a separate network and only shares some physical resources with user plane traffic. It has the disadvantage that it requires a very expensive dedicated network infrastructure.

Recently, many proprietary solutions that allow signalling transport over IP have appeared. This approach enables the use of a common core network for transporting both signalling and media traffic. The Stream Control Transmission Protocol (SCTP) [2] is a standards-based solution to this problem. It is the result of the work of the IETF Signalling Transport (SIGTRAN) group, which was formed in 1999. Our work for this paper concentrates on one of the features of SCTP, which is called “multi-homing”. This gives SCTP a certain amount of network level fault tolerance by using several IP addresses for each single endpoint.

This paper first introduces briefly the SCTP protocol and its “multi-homing” aspect. After describing the experimental set-up, it shows the mechanisms of SCTP through the results of experiments performed in three different situations: “multi-homed” SCTP over a loss-less link, “single-homed” and “multi-homed” SCTP over a lossy link. This paper also describes the future work that will be done to get a better knowledge of the performances of the “multi-homed” feature of SCTP.

### 1.1 Role of SCTP

In creating SCTP, the IETF aimed to develop a new transport protocol that would overcome the limitations of TCP (Transmission Control Protocol) and UDP (User Datagram Protocol) to match the requirements put on a signalling bearer in a telecommunication network.

Therefore, SCTP is designed to be particularly effective for applications that:

- require framing of reliable data streams,
- require ordered transport of data, but can transfer multiple message sequences that are unrelated,
- need to transfer messages that hold no particular sequence or relationship to one another or can be correlated and sequenced at the application level,
- require network layer redundancy (to avoid losing too much efficiency when a fault occurs at the transport layer or below).

UDP is connectionless, which means that it cannot provide error control and flow control. These are vital properties of a transport to support the real-time transfer of signalling information. TCP, even though it is connection oriented, has other drawbacks that decrease its suitability. It is byte stream oriented, always delivers data in strict order, does not allow control of protocol parameters and does not support “multi-homing”.

SCTP operates like TCP (using window-based flow control) with the additional features necessary to transport signalling information. SCTP supports data exchange between two endpoints. It provides reliable transmission, detecting when data is discarded, reordered, duplicated or corrupted and retransmitting damaged data as necessary. SCTP is rate adaptive similar to TCP, and will scale back data transfer to the prevailing load conditions in the network. It is designed to behave co-operatively with TCP sessions attempting to use the same bandwidth. And to provide network layer redundancy, SCTP includes the “multi-homing” capability, which is described in the following section.

## 1.2 SCTP Multi-homing – An Overview

An SCTP multi-homed host has more than one IP address that can be used to reach it. So if the routing is configured in such a way that these IP addresses are accessible through different paths, multi-homing gives SCTP a certain network level fault tolerance. In the case of a failure on the path to one of the IP addresses, the SCTP sender may use another of the receiver’s addresses (example shown in Figure 1).

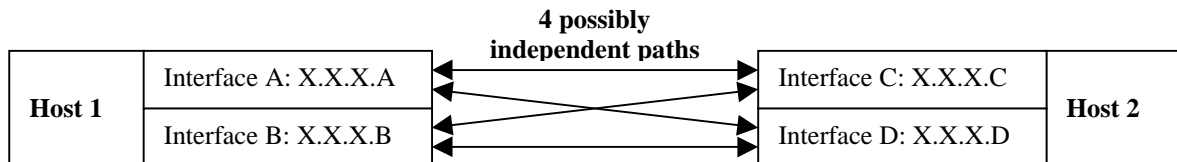


Figure 1: Network level fault tolerance of Multi-homing

Depending on network configuration this may result in a different path being used for communication between the two hosts.

As specified in RFC 2960 [2], SCTP multi-homing support adds the following basic functions to the protocol:

- within an association, one path is considered primary. This means that one of the IP addresses assigned to the receiver endpoint of the association is selected to be the primary address. The primary path corresponds to the network path that leads to the primary address of the peer endpoint. Unless specified otherwise by the SCTP user, an endpoint should always transmit on the primary path to its peer endpoint.
- when acknowledging received chunks, the reply chunks (acknowledgements) should be transmitted through the same path that was used by the received chunks. However, in the case where the receiver acknowledges several data chunks received in packets from different source addresses at once, the acknowledgement chunk may be sent to any of these source addresses.
- when retransmitting a data chunk to a multi-homed peer, the receiver should choose another destination address than the one to which the original data chunk was sent. If a chunk is not acknowledged and has to be retransmitted, there might be a network failure. As well, when an endpoint receives a duplicate data chunk, the endpoint may vary the destination address and not use the source address of this duplicate data chunk to send an acknowledgement.
- another feature of SCTP is used to help the implementation of multi-homing: the heartbeat mechanism. This mechanism detects failures in idle paths (path that are not being used) and endpoints and therefore can detect if a destination address is active or passive. Heartbeat chunks are sent periodically to all idle destinations, and a counter is maintained on the number of Heartbeats sent to an inactive destination without receipt of a corresponding Heartbeat Acknowledgement. When this counter exceeds a configured maximum, that destination address is declared inactive. So when the need is to use another IP address than the primary address to reach a multi-homed endpoint, SCTP knows which other addresses are active or not and thus, can avoid using another path that has a failure.

Thus, the RFC gives great flexibility when implementing this multi-homing feature as it uses “should” or “may” when describing its behaviour. These correspond to IETF keywords for optional parts of a specification.

## 2. Experimental Set-up

### 2.1 The experimental network

To perform our investigation an experimental network was used. This network consists of a Linux router that connects two LANs. The router emulates a WAN, using the NISTNet software [3]. This is similar to an earlier SCTP experimental environment discussed in Jungmaier et.al. [4]. Each of the two LANs is made of a ethernet hub that connects one multi-homed SCTP host to the router. Figure 2 illustrates this network with some example characteristics.

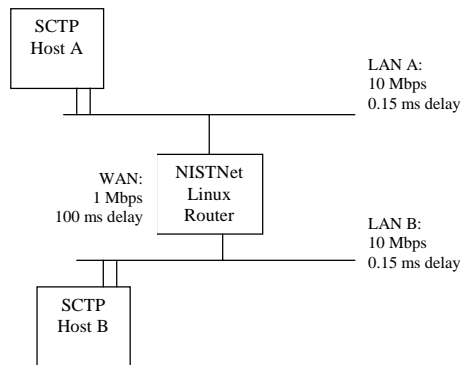


Figure 2: Experimental Set-up

Each SCTP host is realised by using a computer running the Reference Implementation of SCTP [5]. Each SCTP host computer is equipped with two Ethernet cards, to enable multi-homing.

### 2.2 Capabilities of Experimental Set-up

The physical configuration of this set-up makes possible to work with either multi-homed or single-homed SCTP. This gives us the ability to compare these two configurations.

The NISTNet tool emulates a WAN between our two LANs. It provides with the possibility to emulate bandwidth limitations, delays and packet loss independently on each of the routes through this WAN. This feature allows us to have different characteristics between the primary and the secondary paths, as illustrated on the Figure 3.

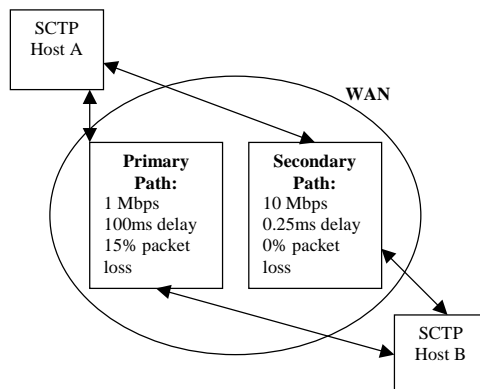


Figure 3: Experimental set-up capabilities

In order to collect data from experiments, it was necessary to modify the SCTP stack. These modifications consisted of adding code to write a number of data files during an experiment. The data collected included the Transmission Sequence Number TSN of each sent packet, the time at which the packet was sent and the interface the packet was sent on. This allows the investigation of the implementation of the SCTP multi-homing feature by recording which packets are sent on the secondary path.

The Transmission Sequence Number corresponds to the number given by SCTP to each data packet sent. This sequence number is used by SCTP to enable congestion control, acknowledgement of data and to preserve packet ordering.

### 3 Results and Analysis

Three experimental scenarios were investigated. These are as follows:

- multi-homed hosts with bandwidth limitation and delay but no packet loss,
- single-homed hosts with bandwidth limitation, delay and packet loss,
- multi-homed hosts with bandwidth limitation, delay and packet loss.

All these three cases are based on the experiment model shown in Figure 3. Each scenario consists of SCTP host A transmitting a stream of 15000 byte data packets to SCTP host B. For scenarios with packet loss, there is a packet loss rate of 15% configured on the router, such that multi-homing is invoked. 15 % is a very high rate for packet loss, but it was needed to trigger the multi-homing mechanism in the context of these experiments. The bandwidth was limited to 1 Mbps and a delay of 100 milliseconds was added through the primary path. Through the secondary path, the bandwidth was limited to 10 Mbps and the delay was set to 0.25 milliseconds.

#### 3.1 Multi-homed with no packet loss

The packet loss was set to 0%. Thus, multi-homing was not triggered and only the primary path was used (case identical to the single-homed case). Figure 4 shows the TSN number of the sent packets versus the time at which they were sent.

The plot in Figure 4 shows the way SCTP normally sends packets, when no loss is involved. It is possible to see that SCTP sends data up to the congestion window size value and then waits for acknowledgements to put more data onto the network. Also, the time between every series of sent packets corresponds to the round trip time, the time it takes the packet to reach its destination and to be acknowledged. This round trip time is equal to the sum of the delays in the network: 0.15 (Packet through the first LAN) + 100 (packet through the WAN) + 0.15 (packet through the second LAN) + 0.15 (acknowledgement through the second LAN) + 100 (ack through the WAN) + 0.15 (ack through the first LAN) = 200.6 ms. This plot also shows the “opening” of the congestion window *cwnd*, which means that every time a packet is acknowledged the congestion window size is growing. *Cwnd* corresponding to the amount of data that the network can handle. This is the “slow-start behaviour”.

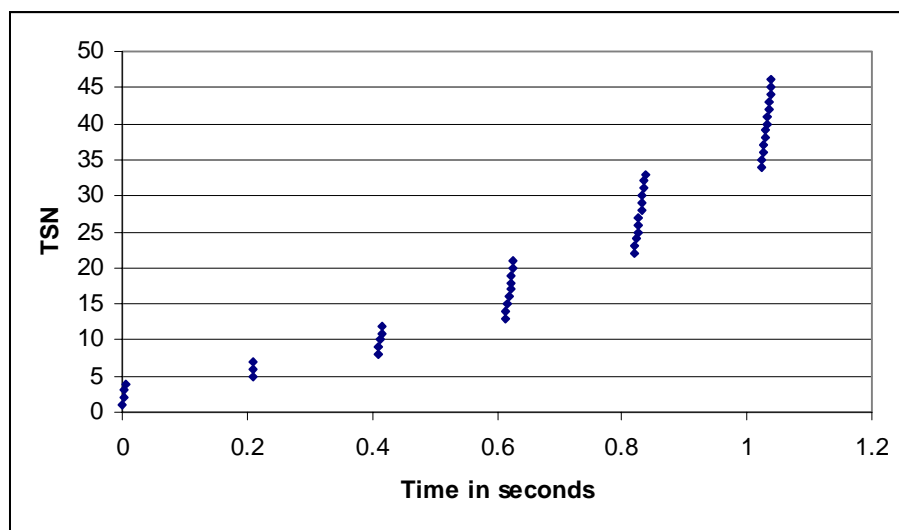


Figure 4: Multi-homed with no packet loss

For deeper explanation, during the “slow start” phase of SCTP *cwnd* increases its value of one packet each time the sender receives an acknowledgement. One acknowledgement being sent every two received packets.

For this plot, *cwnd* starts with the value  $cwnd = 4$ . The second set of transmissions occurs when SCTP has received the ACK generated by the 2<sup>nd</sup> received packet. At this time there is still data in flight in the network and so SCTP can only put new data into the network equal to the data acknowledged plus the increase in *cwnd* for receiving an ACK in slow-start (= 1 MTU, or packet). In other terms, that is why we only send three packets in the second set of transmission.

### 3.2 Single-homed with packet loss

Single-homed means there is only one path to reach the peer endpoint. The bandwidth limit was set to 1 Mb through the Linux router. This had the effect to create a bottleneck, as the LAN bandwidth limit is of 10 Mbps. Also, a delay of 100 ms was set through the same router and so was a packet loss rate of 15%.

Figure 5 shows the TSN plot for this scenario., triangles standing for retransmitted packets.

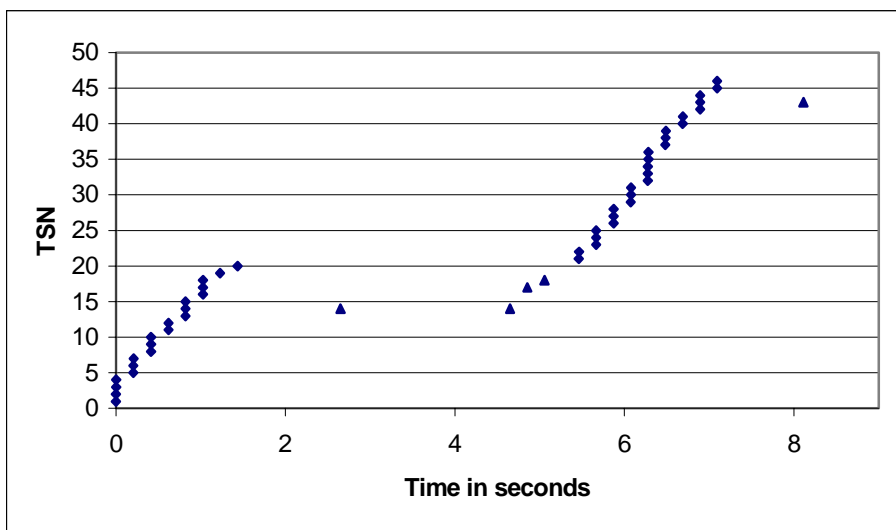


Figure 5: Single-homed with packet loss

This plot shows a slow start phase interrupted by the loss of a first packet. Then, what happens is that the retransmitted packet for the first lost packet is also lost. The effect is very important on the transmission, which is almost stopped for more than 3 seconds here. Then comes the retransmission of two other lost packets. SCTP returns finally to its slow start phase and recovers from these consecutive losses.

So one big danger of retransmitting lost packets onto the same path is that if the packet was lost, there has to be something wrong with the link, which means that the retransmitted packets could go wrong as well. And as the Figure 5 shows it, it takes SCTP a good while to recover from this situation.

### 3.3 Multi-homed with packet loss

The exact same set-up as described in the Figure 3 was used here. The Figure 6 corresponds to the TSN plot of an experiment performed following this scenario.

For this experiment, another information was gathered: which path the SCTP stack used to transmit the packets. The experiments have shown that the secondary path was used for every retransmitted packet and only for these. This means that the Reference Implementation only uses multi-homing in a redundancy purpose and not with the goal of sharing the bandwidth between the two paths. On the plot the circles stand for retransmitted packets, these used the secondary path.

The plot also shows that the retransmitted packets don't get lost on the secondary path, as the loss probability was set to zero on this one. So using the multi-homing feature decreases the risk of losing a retransmitted packet as it is likely that both the primary and secondary path are encountering troubles at the same time.

It is not yet clearly possible to say if the multi-homing configuration recovers quicker from the loss of packets than the single-homed one as the values of the receiver window size and the congestion window size are not available yet for every time SCTP transmits a packet and as the acknowledgement are not yet written on the plots.

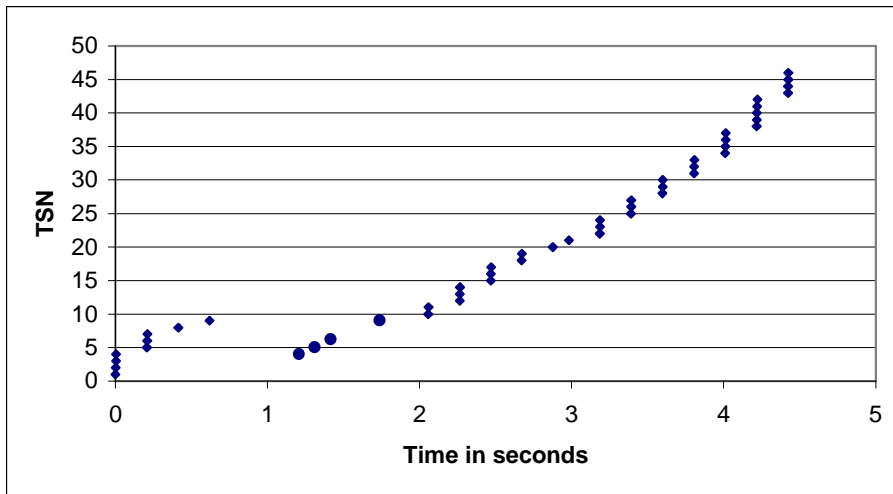


Figure 6: Multi-homed with packet loss

## 4. Future Work

The first thing to be done is to add a few more functions to the SCTP stack to gather more data during the experiments. Being able to plot the acknowledgements of the packets versus the time at which they arrive will allow a better understanding of the TSN plots. Writing the values of the congestion window size *cwnd* and the receiver window size *rwnd* when every packet is sent will also help a lot to understand these plots better. Another important thing to write is the time at which the packets are received at the peer endpoint, and the size of each packet to be able to calculate the throughput. As the tests are performed with a random probability of packet loss, it will be necessary to run these tests a hundred times and calculate a mean for the throughput to obtain a proper result.

When these functions are implemented, the goal will be to expand the tests to different scenarios of experiments. In that way, the effect of different link characteristics will be explored. The effect of the heart-beating mechanism when using multi-homing will also be investigated. As well, the link changeover policy used in the Reference Implementation of SCTP will be investigated to lead to possible improvements. Finally, the accent will be pointed on the SCTP protocol parameters. Varying these parameters may give more effective configuration depending on the type of environment (Internet/Intranet). It may also give us insight into ways in which the link changeover policy in SCTP can be improved.

## Conclusion

The work presented here shows the multi-homing behaviour of the Reference Implementation of SCTP. We use plots that show the TSN of SCTP packets versus the time at which they are sent. These have been made possible by the extra experimental data collection features that we added to the SCTP stack and our experimental environment. This initial work concentrates on investigating how the Reference Implementation of SCTP deals with packet loss and implements multi-homing. This experimental set-up will form the basis of our future work on optimised SCTP link change over policies.

This paper is a first step in evaluating SCTP multi-homing performance. It is planned to perform further experiments that concentrate on throughput measurements for both single-homed and multi-homed SCTP. It is hoped that this work will allow us to make recommendations on suitable SCTP parameter configurations for differing network environments and to compare SCTP and TCP throughput characteristic.

## References

- [1] ITU-T Recommendation Q.700: Introduction to CCITT Signalling System No. 7, International Telecommunication Union, Geneva, March 1993.
- [2] R. Stewart, Q. Xie, et al., "RFC 2960: Stream Control Transmission Protocol", The Internet Society, 2000.
- [3] Carson: NIST Network Emulation Tool, <http://www.antd.nist.gov/itg/nistnet/>, March 2001.
- [4] A. Jungmaier, M. Schopp, M. Tüxen, "Performance Evaluation of the Stream Control Transmission Protocol", Forum on Design Languages (FDL) '99, Ecole Normale Supérieure de Lyon, FRANCE, Aug. 30 - Sept. 3, 1999.
- [5] Available on Randall Stewart's website: <http://www.sctp.org>.