INTERNET-DRAFT
Internet Engineering Task Force
Issued: 8 may 2000
Expires: 30 October 2000

L. Coene
Siemens
J. Loughney
Nokia
I. Rytina
Ericsson
L. Ong
Nortel Networks

Stream Control Transmission Protocol Applicability Statement

<draft-ietf-sigtran-sctp-applicability-01.txt>

Status of this Memo

Abstract

This document describes the applicability of the Stream Control Transmission Protocol for general usage. A few general applications are described such as the transport of signalling information (SS7, DSS1/2 ...) over IP infrastructure. The use and specification of adaptation layers in conjunction with SCTP is described.

1 Introduction

This document covers subject terminology and makes a overview of the solutions for transporting information over Internet Protocol infrastructure. The transport medium used is the Stream Control Transmission Protocol (SCTP). However some of the issues may also relate to the transport of information via TCP.

SCTP provides the following services to its users:

- acknowledged error-free non-duplicated transfer of user data

- transport-level segmentation to conform to discovered MTU size

- sequenced delivery of user datagrams within multiple streams, with an option for order-of-arrival delivery of individual datagrams

- optional multiplexing of user datagrams into SCTP datagrams, subject to MTU size restrictions

- enhanced reliability through support of multi-homing at either or both ends of the association.

- Explicit indication in the message of the application protocol SCTP is carrying.

1.1 Terminology

The following terms are commonly identified in related work:

Port Number: Indicates on the transport level which application needs to be reached in the layer above. Transport Address: An IP address and a port number forms a transport address which identifies a SCTP association. Protocol Identifier: Indicates the upper layer protocol that is using SCTP for the transport of its data. Chunk: a unit of information within an SCTP datagram, consisting of a chunk header and chunk-specific content. Each chunk can contain user or data information about the particular SCTP association. Multihoming: Endpoint which uses more than one IP address for receiving SCTP datagrams on the same association. NAT: Network Address Translation SACK: Selective Acknowledgement message, this is a response on the data msg acknowledging the receipt of it at the remote side. TSN: Transaction Sequence Number, this is a number assigned by SCTP to assure reliable delivery of user data within an association.

2 Stream Control Transmission Protocol -- SCTP

2.1 Introduction

The Stream Control Transmission Protocol (SCTP) provides a high reliable, redundant transport between two endpoints. The interface between SCTP and its applications is handled via adaptation layers which provide a intermediate layer so that the existing upper layer protocols do not have to change their interface towards the transport medium and internal functionality when they start using SCTP instead of an other transport protocol.

The following function are provided by SCTP: - Initialization of transport association - Synchronization of association state - Synchronization of sequence numbering - Reliable Data Transfer - Forward and backward sequence numbering - Timers for transmission and acknowledgement - Notification of out-of-sequence - Retransmission of

lost messages - Support of multiple control streams - Separate sequence control and delivery of each stream - Congestion control - Window flow control - Congestion avoidance based on TCP methods, e.g. using
retransmission backoff, window reduction, etc. - Detection of session failure by active means, e.g. heartbeat - Termination of association SCTP does support a number of functions that are not provided by current TCP: - no head-of-line blocking, i.e. multiple streams - multilink failover for added reliability - keep-alive function for active rapid failure detection - message verses byte sequence numbering - tighter timer control (than standard TCP implementations)

By defining the appropriate User adaptation module, a reliable transport mechanism can be provided: - reliable transmission of packets with end-to-end congestion control provided using methods similar to TCP - choice between sequenced and unsequenced, reliable message delivery - keep-alive message

Within a association between the two endpoints, 1 or more stream(s) may be available. These streams are visible to the adaptation layers but are invisible to any layer above the adaptation layer.

2.2 Issues affecting deployment of SCTP

2.2.1 SCTP Multihoming

Redundant communication between 2 SCTP endpoints is achieved by using multihoming where the endpoint is able to send/receive over more than one IP address.

Under the assumption that every IP address will have a different path towards the remote endpoint, (this is the responsibility of the routing protocols or of manual configuration), if the transport to one of the IP address (= 1 particular path) fails then the traffic can migrate to the other remaining IP address (= other paths) within the SCTP association.

As a practical matter, it is recommended that IP addresses in a multihomed endpoint be assigned IP endpoints from different TLV's to ensure against network failure.

Multihoming provides redundant communication in SCTP by allowing communication between two endpoints to continue in the event of failure along a path between the endpoints.

SCTP will always send its traffic to a certain transport address (= destination address + port number combination) for as long as the transmission is uninterrupted (= primary). The other transport addresses (secondary paths) will act as a backup in case the primary path goes out of service. The changeover between primary are backup

will occur without packet loss and is completely transparent to the application.

The port number is the same for all transport addresses of that specific association.

Applications directly using SCTP may choose to control the multihoming service themselves. The applications have then to supply the specific IP address to SCTP for each datagram. This might be done for reasons of load-sharing and load-balancing across the different paths. This might not be advisable as the throughput of any of the paths is not known in advance and constantly changes due to the actions of other associations and transport protocols along that particular path, would require very tight feedback of each of the paths to the loadsharing functions of the user.

Applications using adaptation layers to run over SCTP do not have that kind of control. The adaptation layers will have to take care of this.

By sending a keep alive message on all the multiple paths that are not used for active transmission of messages across the association, it is possible for SCTP to detect whether one or more paths have failed. SCTP will not use these failed paths when a changeover is required.

The transmission rate of sending keep alive message should be modifiable and the possible loss of keep alive message could be used for the monitoring and measurements of the concerned paths.

2.2.2 Fast retransmit of chunks

The retransmission of a message is basically governed by the retransmission timer. So if no acknowledgement is received after a certain time, then the message is retransmitted. However there is a faster way for retransmitting which is not dependant on that timer.

Every second message that a node received will be acknowledge to the remote peer. If gaps occur in the acknowledge message at the remote side, then the remote side will wait 3 further gap reports(acknowledgements) before it retransmit the message. As the gap occurs, the node must transmit a SACK on every datagram until there are no more gap. This retransmission will happen far sooner than with a timer. Especially if the traffic volume increases in SCTP, those retransmissions of the chunks would happen faster and faster (and hopefully, they would also be faster acknowledged). In any case if gaps occur, the node will certainly try to acknowledge them faster(irespective of the fact if the SACKs will get to the remote node, where, if received, they would speed up the retransmission of the chunks)

See also the paragraph on congestion control and avoidance.

2.2.3 Use of SCTP in Network Address Translator (NAT) Networks

When a NAT is present between two endpoints, the endpoint that is behind the NAT, i.e., one that does not have a publicly available network address, shall take one of the following options:

A) Indicate that only one address can be used by including no transport addresses in the INIT message. This will make the endpoint that receives this Initiation message to consider the sender as only having that one address. This method can be used for a dynamic NAT, but any multi-homing configuration at the endpoint that is behind the NAT will not be visible to its peer, and thus not be taken advantage of.

B) Indicate all of its networks in the Initiation by specifying all the actual IP addresses and ports that the NAT will substitute for the endpoint. This method requires that the endpoint behind the NAT must have pre-knowledge of all the IP addresses and ports that the NAT will assign.

This requires the adaptation of NAT boxes to search within SCTP outgoing INIT and incoming INIT_ACK mesages for the addresses and replace them with the NAT internal address in addition to replacing the addresses in the IP header.

C) Use RSIP [RFCRSIP] where the connection is tunneled from host until the NAT border and the host layers above IP network layer have no knowledge of the NAT internal addresses.

D) Use the hostname feature and the DNS to resolve the addresses. (Ed note: have to figure out hows this precisely works)

2.2.4 MTU path discovery

SCTP discovers the maximal length of the message that can be transported through the network to the final destination without having to fragment(=chop something in pieces) the message in IP network layer. This avoids using IP fragmenting. SCTP level segmentation is beneficial because if a packet is lost during network transmission, only that packet will need to be retransmitted. Contrasted with IP-level segmentation, where the whole unsegmented message will have to be retransmitted, this is a much more effective scheme [RFC1981].

2.2.5 Use of multiple streams

A stream in a one-directional stream of bytes between 2 endpoints within a SCTP association. A association can have one or more streams

in its association and the number of streams in one direction does NOT need to be the same as the number of streams in the opposite direction. The number of streams in both directions is thus assymmetrical.

The application can choose on which stream it can send it data. Streams may specify order of deliver or sequenced delivery. Some application level protocols may reserve certain streams for certain media, for example sending graphical content (jpeg, gif, etc.) of a web page through a certain stream while text through others, and streaming content through others. Any packet loss on one stream will not block packet transmission on others.

Each stream within a association should be looked upon as a link between two points. If multiple streams are used then the application is dealing with multiple links towards the destination. Some applications require the use of sequenced delivery, which would require for them to select a certain link to send their message on.

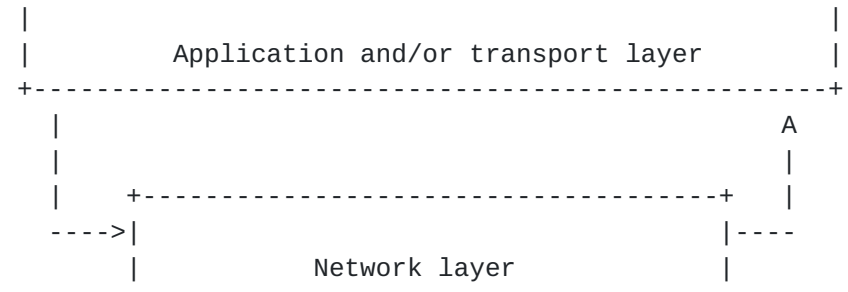
2.2.6 Congestion control & avoidance

Congestion control and/or avoidance is of primordial importance in any connectionless network. Congestion is the result of approaching or exceeding the processing capacity of the link, network, application and/or transport layers. If the processing capacity is exceeded, then the congestion can be avoided (example taking a other non-congested path towards the destination) or controlled (for example, reducing the rate of messages to that destination).

The reaction of SCTP to congestion is detailed in the next paragraphs.

Congestion can be controlled and/or avoided on different levels: - Transport: congestion control/avoidance within SCTP, TCP(fig 2.1.2) - Network : Congestion control/avoidance present in the network layers(example: SCCP, MTP ...) - Link layer: flow control

SCTP conforms to the model of end-to-end congestion control (Fig 2.2.6.2) [RFCSALLY] while ISUP and SCCP model themselves on a link and network based congestion control/overload mechanism (Fig 2.2.6.3).



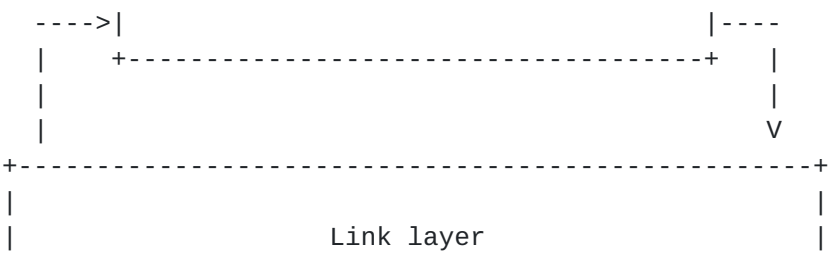


Fig 2.2.6.1 General Congestion model

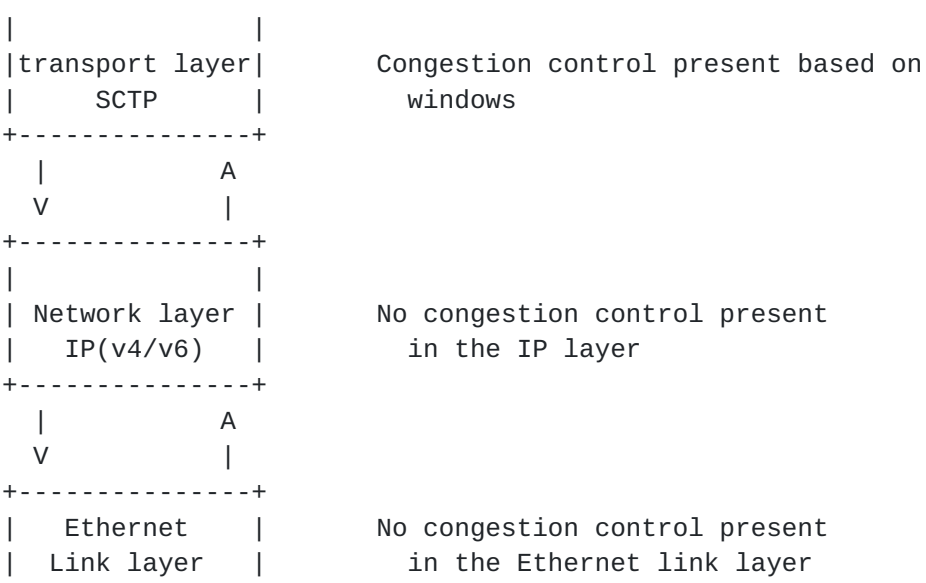


Fig 2.2.6.2 End-to-End congestion control

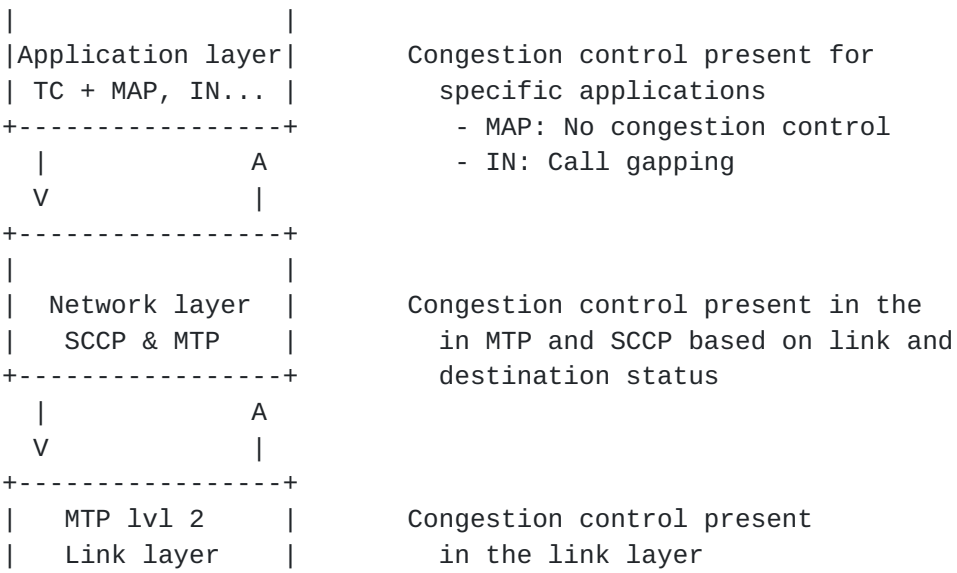


Fig 2.2.6.3 Distributed congestion control

By default, SCTP associations do not have a fixed capacity assigned to them unless other QoS mechanisms are employed. Thus congestion within SCTP association can and will be affected by all traffic using the same links including other SCTP, TCP, RTP, UDP ... traffic traveling on the same path followed by the SCTP association.

2.2.6.1 3-SACK rule in SCTP.

The Selective Acknowledgement (SACK) is one of the cornerstones of SCTP. It selectively Acknowledges datagrams that have been successfully received by the remote node. It serves 2 purposes: - it indicates until a certain datagram that all previous datagrams have been received (without any holes in the sequence) and - it indicates the datagrams sequence ranges which have been received (and so does indicate the holes/gaps between them). It provides us with a form of gap/hole report on messages that have been lost or delayed. A hole can consist of one or more messages.

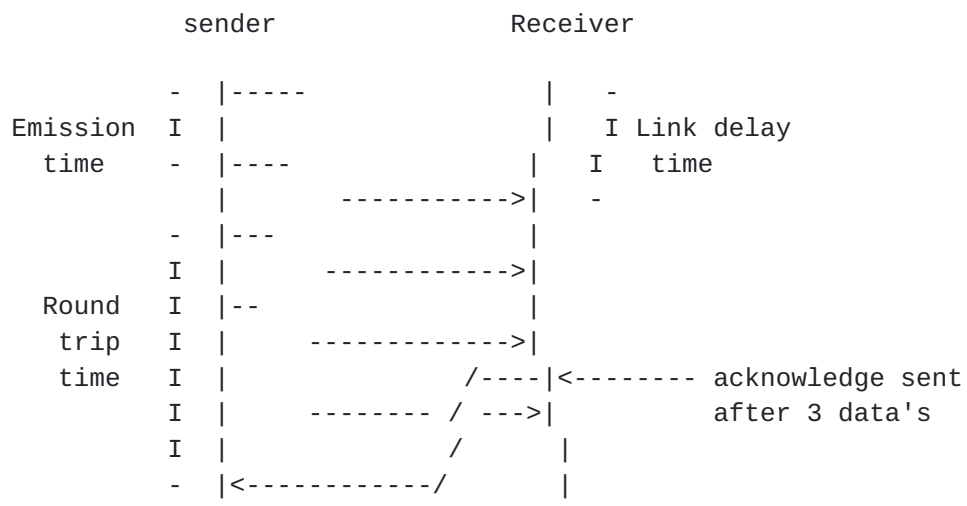


Fig 2.2.6.4 Influence of Window Size/ Link Speed/ Round Trip Delay

Fig 2.2.6.4 is given here as a example where after receiving 3 messages an advisory acknowledgement (SACK) is sent (in this case window = 6). Therefore the sender could be kept busy. The acknowledgement opens the window again. The total time (from first emission till the receiving of the acknowledgement) calculates as: (max. windowsize * emission time)/2 + round trip delay. If the round trip time(RTT) is large, the advisory acknowledgments (SACK) will enhance the throughput.

The SACK is always generated and send back to the sender either - after every second message received (delayed ack). - after at most 200ms after receiving the last message.

The reason for the holes may be diverse: - simple message loss - different round trip times of messages being transmitted on different interfaces

At the sender end, whenever the sender notices a hole in a SACK, it should wait for 3 further SACKs (identifying the same hole) before taking action. This is 3 strikes besides the first one, so that means 4. Thus after 4 SACK, the datagrams belonging to the hole should be retransmitted(and only those).

If gaps occur, the receiver end will send SACKs on every data message received instead on being send on every second data message received. As the sender is waiting for the 3 SACK strikes and the receiver is increasing the SACK rate, that would mean that retransmission would be happening faster. Also the window should be opening up more than in the normal case (= transmission without gaps).

The 3 SACKs rule might be relaxed in certain networks provided certain condition are met:

- private IP network - closed networks - only a single type of application traffic is running on that network (the message in the network exhibit the same characteristics:
 example: signalling messages).

The SACK rule might be configurable in such a networks, if the network operator felt confident in the correctness of the network. This would mean that in case of packet loss, retransmission could be "immediate".

SACK will also report duplicate message arrival. See paragraph 2.2.6.4.

2.2.6.2 Congestion Control

The number of messages in flight is determined by the Congestion window (Cwnd). Every time a message is SACK, a new message might be send to the remote side(up till the Cwnd), even if gaps exists which might ultimately lead to retransmissions.

The value of the Cwnd is dependant on the slow start and/or congestion avoidance/control.

If messages are getting lost, then it is assumed by SCTP that they are lost according to congestion, not that they are lost due to error on the link(such as cable cutthrough ...).

When messages are lost then the rate of messages sending is reduced, till no messages are lost.

2.2.6.3 Use of Explicit Congestion notification (ECN)

Explicit Congestion control is a experimental method for communicating congestion back to the end node. SCTP does not support the use of ECN, but specific recommendations for using ECN with SCTP might be forthcoming.

2.2.6.4 Duplicated messages

SACKs can get lost. The receiving node would then received duplicated packets. A reason for such a behavior is imbalance between the 2 traffic direction, use of different up and down path.

(Ed note: something more has to be put here, still thinking on the right words and reading a couple of RFCs on the subject :-)

2.2.6.5 SCTP in high throughput delivery networks

The TSN is associated with a message, not with the number of bytes(as is the sequence number of TCP) in the message. So the TSN will wrap around less frequently but has a dependency on the length of each message. Use of short messages will lead to a faster wrapping around of the TSN. So in high throughput networks, it is advised to make the messages as long as possible so that the wrap around will be less frequent.

SCTP already has a larger window than TCP does even when TCP is using the "large windows" option.

2.2.6.6 SCTP in long delay/Fat networks (LFN)

Long delay(Fat) Networks consists of network paths which have a high "bandwidth*delay product"(such as satelite links(high delay) or high capacity fiber(high bandwith)). There the 3-SACK rule would lead to enhanced throughput, if the initial windowsize is set higher than 2(which is the default value for non-LFNs).

The initial windowsize should be set to a higher value (4 or 8) as that would mean that 4 messages would be injected in the network and the first sack would come back at about the same time as the last message before the window is full, is injected.

Thus to have the most of the 3 sack rule immediatly, the initial window size should at least be set at 4 (and possible at 8 if we are dealing with really very long delays).

The drawback of this is that it makes SCTP more aggressive to begin with(certainly when faced with TCP).

For a more precise description of the issues associated with this, refer to [RFC123], [RFC2001] and [RFC2018]

2.2.6.7 SCTP in Long Thin Networks(LTN)

Long thin networks consists of network paths that traverse "very low bit-rate" links(such as 56 Kbit modem links). This means that a single host can very easy saturate such a link(= pushing the link into congestion).

2.2.7 Use of the protocol identifier in SCTP

Indicates the the upper layer protocol that is using the associations. The protocol identifier is available to the application and is included in each chunk. 0 is the unknown protocol. This protocol id can be used by firewalls for filtering out certain protocols. If firewalls drops certain protocol id then then association will fail in the end because the TSN will be lost. If the chunk(without its user data) is simulated with the TSN in it, then the user data will be dropped, but the association is preserved.

The protocol identifier is administered by IANA[IANA].

2.2.8 Use of QoS methods

SCTP is a end-to-end protocol which cannot guarantee the quality-of-service along the complete path(s) taken by the messages of that particular association. If more guarantees are required for improving the reliability of the transport, some form of QoS mechanism may be needed.

The possible schemes are as follows.

2.2.8.1 Over-provisioning

Over-provisioning of the links so that the total traffic running over the link never exceeds the link capacity. In practice, this may be difficult to ensure reliably.

2.2.8.2 Private Internets

Use of a private network solely for transport purposes. Private networks may allow better control and monitoring of resources available.

2.2.8.3 Differentiated services

By providing a certain code point in the Type-of-service field (TOS), certain Differential services can be selected. [RFC2597, RFC2598]

Setting the code point for transport requires some thought. It is dependant on the kind of differentiate service selected. Also the use

of traffic is important: example signalling info should have a higher priority than the user data traffic for which the signalling is responsible(and that relation does not always exist).

2.2.8.4 Integrated services

By use of integrated services [RFC2208], resources are reserved for signaling transport.

If resources are unavailable for to initiate a new signaling transport, that request will be denied. RSVP may not scale well and this solution may prove to be unfeasible.

An example is Multi Protocol Label Switching.

2.2.9 SCTP Checksum

SCTP uses the Adler-32 checksum algorithm. This algorithm will perform better than a 16 bit (CRC or not) checksum or even a 32 bit CRC checksum.

The message can also be protected by IPSEC which is much stronger. In that case, the checksum should still be computed.

2.2.10 Tunneling of SCTP association over UDP

The basic operation of SCTP is to run directly on top of IP. However, due to restrictions placed on implementers by Operating Systems, not all implementations may be able to run over IP directly. Therefore an alternative is given which might circumvent some or all of the restrictions.

The STCP messages are transported over UDP instead. The following issues must be observed: - the port number in the UDP header should be the port number assigned to SCTP. The port number in the SCTP common header should be the one assigned to the user adaptation layer or to the application of SCTP. This means that port numbers previously used in UDP and/or TCP can be reused for the same application using SCTP. SCTP DOES NOT change the semantics of the port number just because the protocol identifier is added to the SCTP message. - the checksum field might be used as a additional guard against errors(particular errors in the UDP header). However, the SCTP checksum employed is far better at catching errors, but does not take the UDP header into account.

2.2.11 How to define and Use adaptation layers

Many different applications may use SCTP for different purposes. They

go from File transfer over HTTP transport to signalling information transport.

Some applications might want preserve the existing interface with its lower layer (in this case SCTP) while for other applications, this does not pose a problem. A narchitecture has been devised to let the application choose whether they want to run over SCTP directly (just a many applications run over TCP) or let application run on top of a adaptation layer over SCTP.

The basic architecture is as in Figure 2.11.1 :

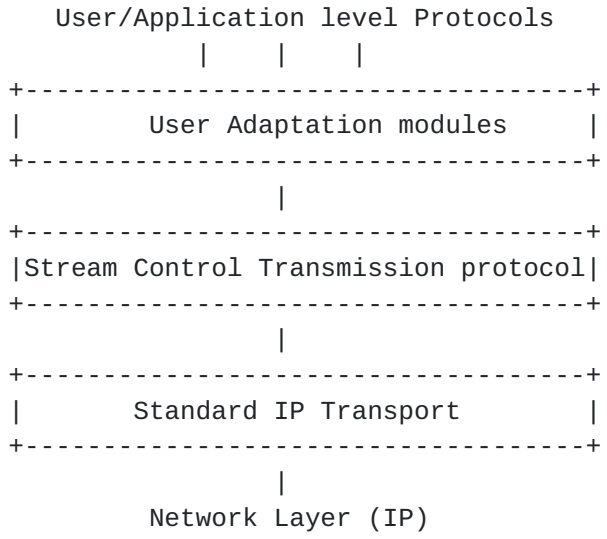


Figure 2.11.1: Transport Components

The three components of the transport protocol are : Adaptation modules that support specific primitives, e.g. management indications, required by a particular user/ application protocol. The use of a adaptation protocol is optional. It is only used in case in which the application protocol does not want to change its interface with the underlying layer.

the Stream Control Transmission Protocol itself that supports a common set of reliable transport functions.

a standard IP transport/network protocol provided by the operating system. In some network scenarios, it has been recognized that TCP can provide limited (but sufficient) reliable transport functionality for some applications.

2.2.12 Security considerations

The following aspects of security are :

Authentication:

Information is sent/received from a known and/or trusted partner.

Integrity:

Information may not be modified while in transit. The integrity of a message in a public network is not guaranteed.

Confidentiality:

Confidentiality of the user data must be ensured. User data can not be examined by unauthorized users.

Availability:

The communicating endpoint must remain in service in all circumstances. Some services have very high availability requirements: for example, all SS7 nodes have to remain active for the 99.999% of the time.

2.2.12.1 General Considerations

SCTP only tries to increase the availability of a network. SCTP does not contain any protocol elements in its messages which are directly related to Authentication, Integrity and Confidentiality functions. It depends for such a features on the IPSEC protocols and architecture.

The only function which has some bearing on security of SCTP is the integrity of message in SCTP, which is guarded by a Checksum. This checksum is mandatory if IPSEC is NOT used. If IPSEC is used then the SCTP checksum becomes optional. The use of IPSEC in the SCTP association must in this case be END-TO-END. The use of IPSEC on a part of a path of a SCTP association does NOT relieve SCTP from using the checksum(as this ain't end-to-end transport)

The general rule is that IPSEC should be turned on unconditionally.

The description of the internet security architecture and the use of it is described in [RFC2401].

2.2.12.2 The cookie mechanism and Denial-of-Service (DOS) attacks

The cookie mechanism in SCTP is a measure against Denial-of-Service (DOS) attacks. In a DOS attack, a lot of init chunks are send towards a single terminating node (the source is a bogus node = a invalid source address in the datagram), so that very quickly all resources are used up and that normal users are rejected due to resource

shortage.

When a INIT chunk is received, the TCB info is encoded and put into the cookie and send to the initiating node via the INIT_ACK. No TCB is allocated at the receiving node as all info is encoded in the cookie and the cookie will return in the COOKIE_ACK (at that time the TCB will be really allocated with the info from the cookie and a full association is set up). As the INIT_ACK will be send back to a bogus address, no COOKIE_ACK will come back and no resources will be tied up in the terminating node.

2.2.12.3 Initiate Tag considerations

As the tag is fixed during the whole lifetime of the association, the initiate Tag values should be selected as random as possible to help protect against "man in the middle" and "sequence number" attacks. It is suggested that RFC 1750 [RFC1750] be used for the Tag randomization. A new tag is only assigned if a new association is set up.

2.2.12.4 Fingerprinting of SCTP

Different implementations may show a certain fingerprint in their messages when they have to answer to certain messages send to them. It is advisabel to send only the basic required information back according to the SCTP protocol.

2.2.12.5 The ACK-Splitting attack

(Ed note : something need to be provided here)

3 Recommendations

To be provided.

4 Adaptation Layers

Currently, there are four adaptation layers, to support carrying of SS7 application protocols over IP. These adaptation layers are being developed for different purposes, and there is no assumption that they should interwork - i.e. - M2UA carries M3UA. They should be thought of as individual protocols for specific uses.

4.1 IUA

There is a need for Switched Circuit Network (SCN) signaling protocol delivery from an ISDN Signaling Gateway (SG) to a Media Gateway Controller (MGC). The delivery mechanism should meet the following criteria

* Support for transport of the Q.921 / Q.931 boundary primitives * Support for communication between Layer Management modules on SG and MGC * Support for management of active associations between SG and MGC

This draft supports both ISDN Primary Rate Access (PRA) as well as Basic Rate Access (BRA) including the support for both point-to-point mode and point-to-multipoint modes of communication. QSIG adaptation layer requirements do not differ from Q.931 adaptation layer, hence the procedures described in this draft are also applicable to QSIG adaptation layer.

4.2 M2UA

There is a need for SCN signaling protocol delivery from an Signaling Gateway (SG) to a Media Gateway Controller (MGC) or IP Signaling Point (IPSP). The delivery mechanism should meet the following criteria:

* Support for MTP Level 2 / MTP Level 3 interface boundary * Support for communication between Layer Management modules on SG and MGC * Support for management of active associations between the SG and MGC

In other words, the Signaling Gateway will transport MTP Level 3 messages to a Media Gateway Controller (MGC) or IP Signaling Point (IPSP). In the case of delivery from an SG to an IPSP, the SG and IPSP function as traditional SS7 nodes using the IP network as a new type of SS7 link. This allows for full MTP Level 3 message handling and network management capabilities.

4.3 M3UA

There is a need for SCN signaling protocol delivery from an SS7 Signaling Gateway (SG) to a Media Gateway Controller (MGC) or IP-resident Database as described in the Framework Architecture for Signalling Transport [11]. The delivery mechanism should meet the following criteria:

* Support for transfer of all SS7 MTP3-User Part messages (e.g., ISUP, SCCP, TUP, etc.) * Support for the seamless operation of MTP3-User protocol peers * Support for the management of SCTP transport associations and traffic between an SG and one or more MGCs or IP-resident Databases * Support for MGC or IP-resident Database failover and loadsharing * Support for the asynchronous reporting of status changes to management

In simplistic terms, the SG will terminate SS7 MTP2 and MTP3 protocols and deliver ISUP, SCCP and/or any other MTP3-User protocol messages over SCTP transport associations to MTP3-User peers in MGCs or IP-resident Databases.

4.4 SUA

This document details the delivery of SCCP-user messages (MAP & CAP over TCAP, RANAP, etc.) over IP. The architecture may be from from an SS7 Signaling Gateway (SG) to an IP-based signaling node (such as an IP-resident Database) as described in the Framework Architecture for Signaling Transport [RFC2719], or between two endpoints located completely within an IP network. The delivery mechanism SHOULD meet the following criteria:

* Support for transfer of SS7 SCCP-User Part messages (e.g., TCAP, RANAP, etc.) * Support for SCCP connectionless service. * Support for SCCP connection oriented service. * Support for the seamless operation of SCCP-User protocol peers * Support for the management of SCTP transport associations between an SG and one or more IP-based signaling nodes). * Support for distributed IP-based signaling nodes. * Support for the asynchronous reporting of status changes to management

5 References and related work

[SCTP] Stewart, R. R., Xie, Q., Morneault, K., Sharp, C. , , Schwarzbauer, H. J., Taylor, T., Rytina, I., Kalla, M., Zhang, L. and Paxson, V."Stream Control Transmission Protocol", <draft-ietf-sigtran-sctp-09.txt>, April 2000. Work In Progress.

[Q1400] SG11, ITU-T Recommendation Q.1400, " architecture framework for the development of signaling and OA&M protocols using OSI concepts ",1993

[HUITEM] Huitema, C., "Routing in the Internet", Prentice-Hall, 1995.

[RFC2373] Hinden, R. and Deering, S., "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[RFC2460] Hinden, R. and Deering, S., "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC814] Clark, D.D., "Names, addresses, ports and routes", RFC 0814, July 1982.

[RFC2401] Kent, S., and Atkinson, R., "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[RFC1981] McCann, J., Deering, S., and Mogul, J., "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[RFC2208] Mankin, A. Ed., Baker, F., , Braden, B., Bradner, S., O`Dell, M., Romanow, A., Weinrib, A. and Zhang, L., "Resource ReSerVation Protocol (RSVP) -- Version 1 Applicability Statement Some

Guidelines on Deployment" , RFC 2208, September 1997.

[RFC2597] Heinanen, J., Baker, F., Weiss, W. and Wroclawski, J., "Assured Forwarding PHB Group", RFC2597, June 1999

[RFC2598] Jacobson, V., Nichols, K. and Poduri, K., "An Expedited Forwarding PHB", RFC2598, June 1999

[RFC2719] Ong, L., Rytina, I., Garcia, M., Schwarzbauer, H., Coene, L., Lin, H., Juhasz, I., Holdrege, M., Sharp, C., "Framework Architecture for Signaling Transport", RFC2719, October 1999

[IANA] Internet Assigned Numbers Authority, http://www.iana.org/, April 2000

[RFCRSIP] Borella, M., Grabelsky, D., Lo, J., Tuniguchi, K. "Realm specific IP",RFCxxxx, xxxx 2000

[RFCSALLY] Floyd, S. Ed., "Congestion Control Principles", <draft-floyd-cong-02.txt> RFCxxxx, April 2000

[RFC1750] Eastlake, 3rd, D., Crocker, S., Schiller, J., "Randomness Recommendations for Security", RFC1750, December 1994

[RFC1323] Jacobson, V., Braden, R., Borman, D., "TCP Extensions for High Performance", RFC1323, May 1992

[RFC2001] Stevens, W., "TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms ", RFC2001, Januarey 1997

[RFC2018] Mathis, M., Mahdavi, J., Floyd, S., Romanow, A., "TCP Selective Acknowledgement Options ", RFC2018, October 1996

6 Acknowledgments

The authors wish to thank Renee Revis, R.R. Stewart, Q. Xie, H.J. Schwarzbauer, M. Tuexen, J.P. Martin-Flatin and many others for their invaluable comments.

7 Author's Address

Lode Coene
Siemens Atea
Atealaan 34
B-2200 Herentals
Belgium

Phone: +32-14-252081
EMail: lode.coene@siemens.atea.be

John Loughney
Nokia Research Center

Itamerenkatu 11-13
FIN-00180 Helsinki
Finland

Phone: +358-9-43761
EMail: john.loughney@nokia.com

Ian Rytina
Ericsson Australia
37/360 Elizabeth Street
Melbourne, Victoria 3000
Australia

Phone : -
EMail:ian.rytina@ericsson.com

Lyndon Ong
Nortel Networks
4401 Great America Parkway
Santa Clara, CA 95054
USA

Phone: -
EMail: long@nortelnetworks.com

Expires: October 30, 2000