Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 15, 2009

R. Stewart
The Resource Group
M. Tuexen
I. Ruengeler
Muenster Univ. of Applied Sciences
July 14, 2008

Stream Control Transmission Protocol (SCTP) Network Address Translation

draft-stewart-behave-sctpnat-04.txt

Status of this Memo

Abstract

Stream Control Transmission Protocol [RFC4960] provides a reliable communications channel between two end-hosts in many ways similar to TCP [RFC0793]. With the widespread deployment of Network Address Translators (NAT), specialized code has been added to NAT for TCP that allows multiple hosts to reside behind a NAT and yet use only a single globally unique IPv4 address, even when two hosts (behind the NAT) choose the same port numbers for their connection. This additional code is sometimes classified as Network Address and Port Translation or NAPT. To date, specialized code for SCTP has NOT yet been added to most NAT's so that only pure NAT is available. The end

result of this is that only one SCTP capable host can be behind a NAT.

This document describes an SCTP specific variant of NAT which provides similar features of NAPT in the single point and multi-point traversal scenario.

Table of Contents

1. Introduction

Stream Control Transmission Protocol [RFC4960] provides a reliable communications channel between two end-hosts in many ways similar to TCP [RFC0793]. With the widespread deployment of Network Address Translators (NAT), specialized code has been added to NAT for TCP that allows multiple hosts to reside behind a NAT and yet use only a single globally unique IPv4 address, even when both hosts (behind the NAT) choose the same port numbers for their connection. This additional code is sometimes classified as Network Address and Port Translation or NAPT. To date, specialized code for SCTP has NOT yet been added to most NAT's so that only true NAT is available. The end result of this is that only one SCTP capable host can be behind a NAT.

This document proposes an SCTP specific variant NAT that provides the NAPT functionality without changing SCTP port numbers. The authors feel it is possible and desirable to make these changes for a number of reasons.

- It is desirable for SCTP end-hosts on multiple platforms to be able to share a global IP address behind a NAT, much as TCP does today.

- If a NAT does not need to change any data within an SCTP packet it will reduce the processing burden of NAT'ing SCTP by NOT needing to execute the CRC32c checksum required by SCTP.

- Not having to touch the IP payload makes the processing of ICMP messages in NATs easier.

2. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Terminology

For this discussion we will use several terms. For clarity we will first define these terms.

- Global-Address - That address that a host behind a NAT is attempting to contact.

- Global-Port - The port number of the peer process at the Global-Address.

- Local-Address - The local address that is known to the host behind the NAT, aka a private address [RFC1918].

- Local-Port - The port number that is in use by the host holding the Local-Address. Normally this is the port that will be translated by the NAPT to a different port number.

- Nat-Global-Address - The global address assigned to the NAT box which it uses as a source address when sending packets towards the Global-Address.

- Natted-Port - The port number that the NAT is using to represent the Local-Port when send data packets toward the Global-Address and Global-Port.

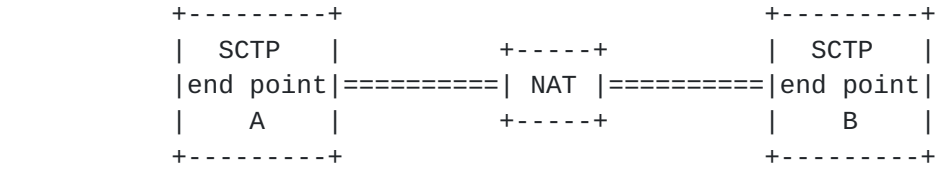
- Local-Vtag - The Verification Tag that the host inside the natted address space has chosen for its communication. The V-Tag is a unique 32 bit tag that must accompany any incoming SCTP packet for this association to the Local-Address.

- Remote-Vtag - The Verification Tag that the host holding the Global-Address has chosen for its communication. The V-Tag is a unique 32 bit tag that must accompany any incoming SCTP packet for this association to the Global-Address.

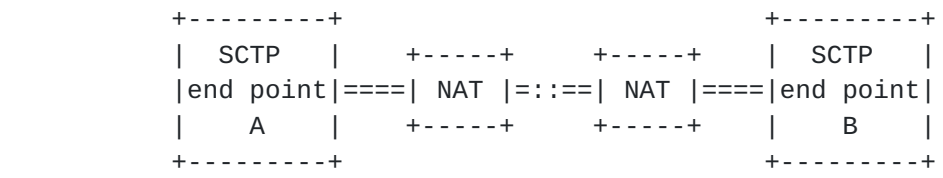
4. SCTP NAT Traversal Scenarios

4.1. Single Point Traversal

In this case, all packets in the SCTP association go through a single NAT, as shown below:



A variation of this case is shown below, i.e., multiple NATs in a single path:

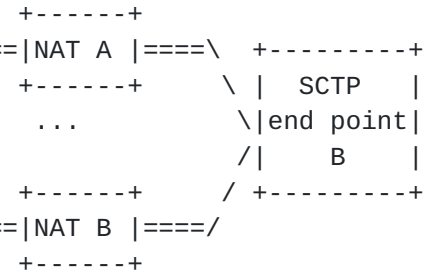


The two SCTP endpoints in this case can be either single-homed or multi-homed. However, the important thing is that the NAT (or NATs) in this case sees ALL the packets of the SCTP association.

In this single traverse point scenario, we must acknowledge that while one of the main benefits of SCTP multi-homing is redundant paths, the NAT function represents a single point of failure in the path of the SCTP multi-home association. However, the rest of the path may still benefit from path diversity provided by SCTP multi-homing.

4.2. Multi Point Traversal

This case involves multiple NATs and each NAT only sees some of the packets in the SCTP association. An example is shown below:



This case does NOT apply to a singly-homed SCTP association (i.e., BOTH endpoints in the association use only one IP address). The advantage here is that the existence of multiple NAT traverse points can preserve the path diversity of a multi-homed association for the entire path. This in turn can improve the robustness of the communication.

To make this work, however, all the NATs involved must recognize the packets they see as belonging to the same SCTP association and perform address translation in a consistent way. It may be required that a pre-defined table of ports and addresses would be shared between the NAT's. Other external management schemes that help multiple NAT's coordinate a multi-homed SCTP association could be investigated.

5. The SCTP specific variant of NAT

In this section we assume that we have multiple SCTP capable hosts behind a NAT which has one Nat-Global address. Furthermore we are focusing in this section on the single point traversal scenario.

The modification of SCTP packets sent to the public Internet is easy. The source address of the packet has to be replaced with the Nat-Global-Address. It may also be necessary to establish some state in the NAT box to handle incoming packets, which is discussed later.

For SCTP packets coming from the public Internet the destination address of the packets has to be replaced with the Local-Address of the host the packet has to be delivered to. The lookup of the Local-Address is based on the Global-VTag, Global-Port, Global-Address, Local-Vtag and the Local-Port.

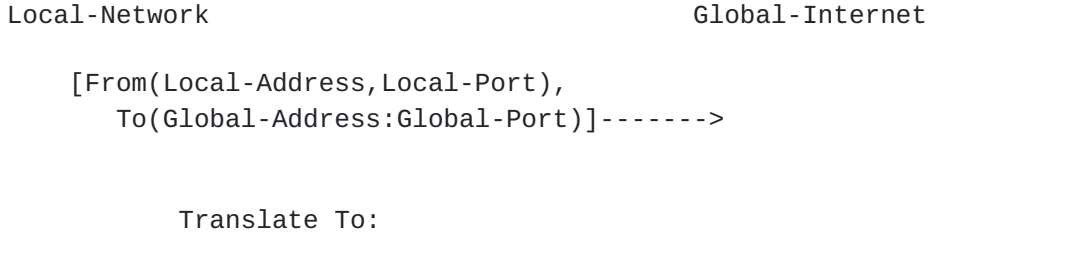
For the SCTP NAT processing the NAT box has to maintain a table of Global-VTag, Global-Port, Global-Address, Local-VTag, Local-Port and Local-Address. An entry in that table is called a NAT state control block.

The processing of outgoing SCTP packets containing an INIT-chunk is described in the following figure.

It should be noted that normally no NAT control block will be created. However it is possible that that there is already a NAT control block with the same Global-Port, Global-Address, Initiate-Tag, Local-VTag but different Local-Address. In this case the INIT SHOULD be dropped and an ABORT MAY be sent back.

The processing of outgoing SCTP packets containing no INIT-chunk is

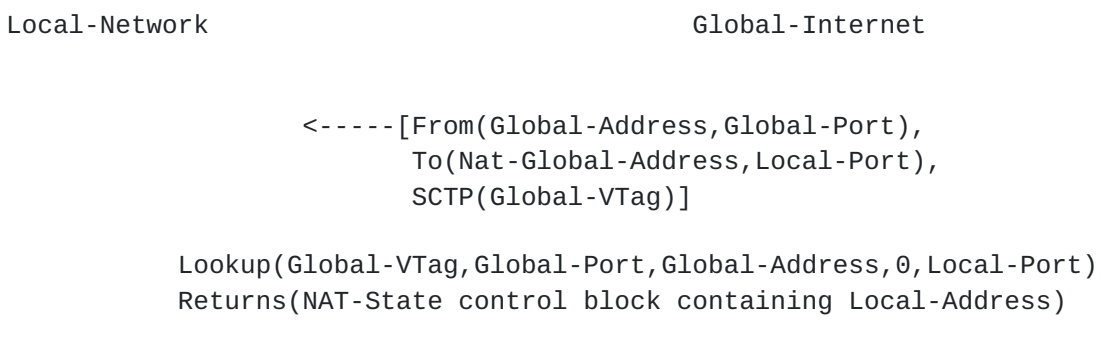
described in the following figure.



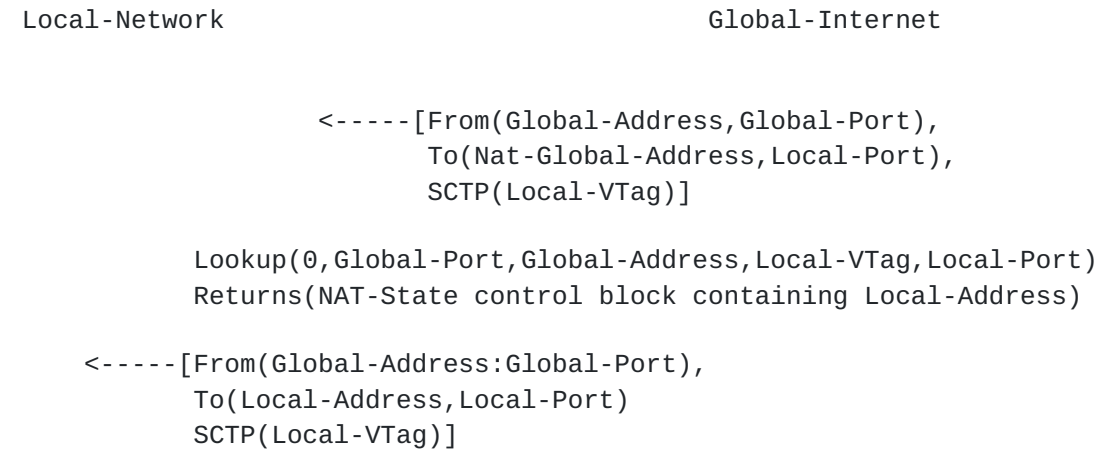
The processing of incoming SCTP packets containing INIT-ACK chunks is described in the following figure.

In the case Lookup fails, the SCTP packet is dropped. The Update routine inserts the Global-VTag in the NAT state control block.

The processing of incoming SCTP packets containing and ABORT or SHUTDOWN-COMLETE chunk with the T-Bit set is described in the following figure.



The processing of other incoming SCTP packets is described in the following figure.



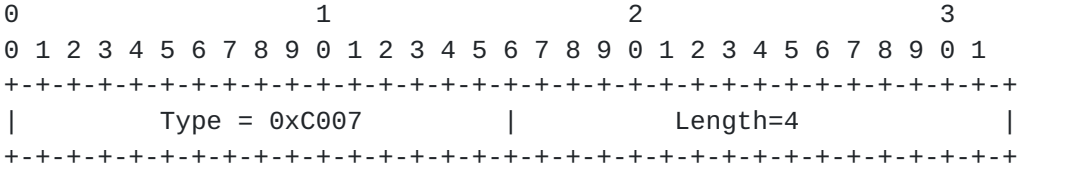
For an incoming packet containing an INIT-chunk a table lookup is made only based on the addresses and port numbers. If an entry with a local vtag of zero is found, it is considered a match and the local v-tag is updated.

This allows the handling of INIT-collision through NAT.

6. Handling of local port number collisions

There is one drawback of the SCTP specific variant of NAT compared to a NAPT solution like the ones available for TCP. Consider the case where two hosts in the Local-Address space want to setup an SCTP association with the same server running on the same host in the Internet. This means that the Global-Port and the Global-Address are

the same. If they both chose the same Local-Port the server can not distinguish both associations based on the address and port numbers. For the server it looks like the association is being restarted. To overcome this limitation the client sends a NAT_SUPPORTED parameter in the INIT-chunk which is defined as follows:



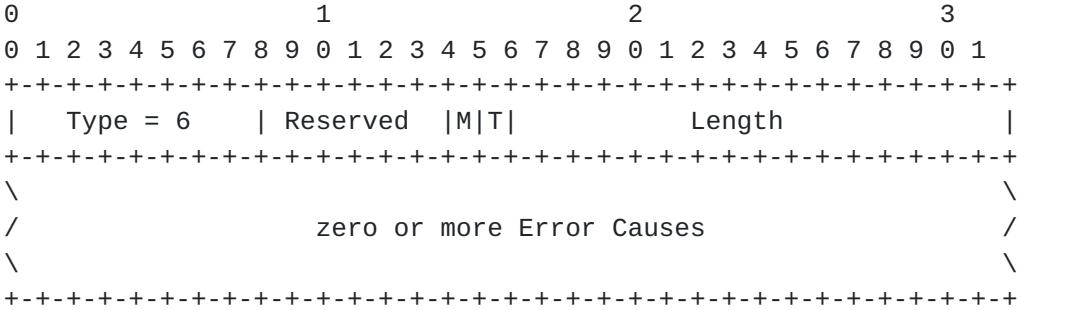
When the server receives this parameter it will also use the verification tag to look up the association. However, this will make it impossible to restart such associations.

7. Handling of local port number and verification tag collisions

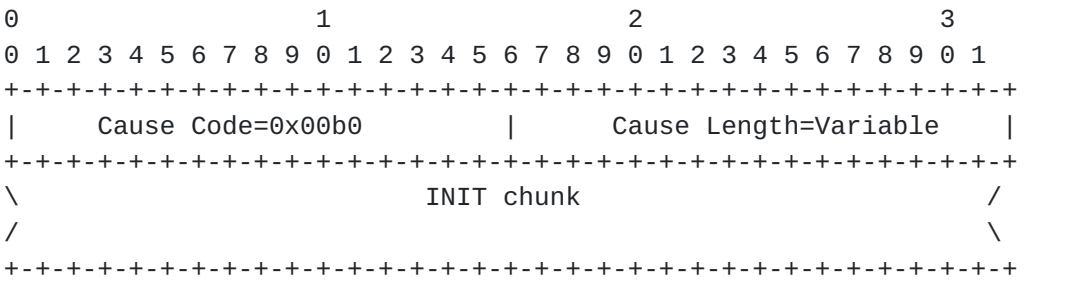
Consider the case where two hosts in the Local-Address space want to setup an SCTP association with the same server running on the same host in the Internet. This means that the Global-Port and the Global-Address are the same. If they both chose the same Local-Port and Local-VTag, the NAT box can not distinguish incoming packets anymore. But this is very unlikely. The Local-Vtags are chosen by random and if the Local-Ports are also chosen ephemeral an random this gives a 46 bit random number which has to match. In the TCP like NAPT case the NAT box can control the 16 bit Natted Port.

However, if this unlikely happens the NAT box MUST respond to the INIT chunk by sending an ABORT chunk with the M-bit set. The source address of the packet containing the ABORT chunk MUST be the destination address of the SCTP packet containing the INIT chunk. The sender of the packet containing the INIT chunk MAY start the association setup procedure after choosing a new initiate tag.

The ABORT chunk defined in [RFC4960] is therefore extended by using the following format:



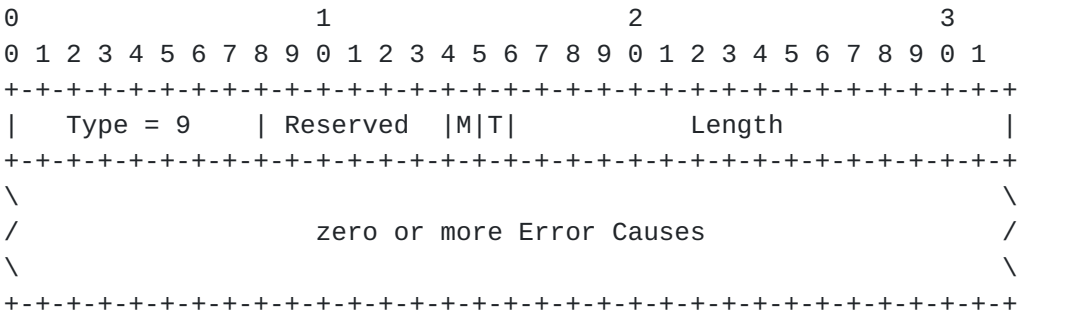
The following error cause with cause code 0x00b0 (Colliding NAT table entry) SHOULD be included in the ABORT chunk:



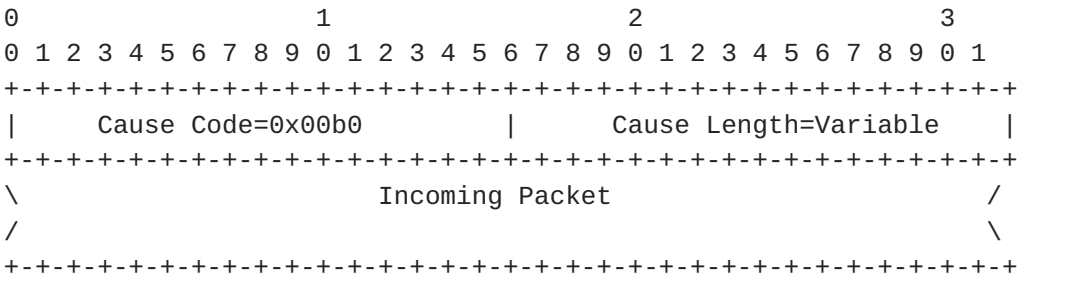
8. Handling of missing state

If the NAT box receives a packet for which the lookup procedure does not find an entry in the NAT table, a packet containing an ERROR packet is sent back with the M-bit set. The source address of the packet containing the ERROR chunk MUST be the destination address of the incoming SCTP packet. The verification tag is reflected.

The ERROR chunk defined in [RFC4960] is therefore extended by using the following format:

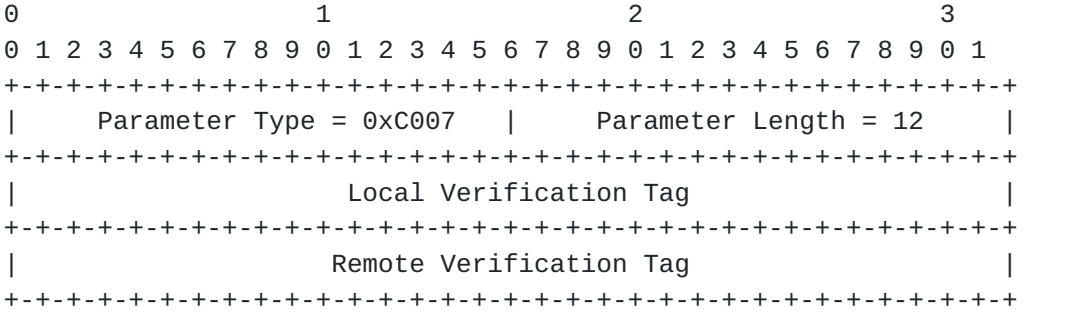


The following error cause with cause code 0x00b1 (Missing NAT table entry) SHOULD be included in the ERROR chunk:



If an end-point receives a packet with this ERROR chunk it MAY send an SCTP packet with a ASCONF chunk containing an Add IP Address

parameter followed by a vtag parameter:



If the NAT box receives a packet for which it has no NAT table entry and the packet contains an ASCONF chunk with a vtag parameter, the NAT box MUST update its NAT table according to the verification tags in the vtag parameter.

9. Multi Point Traversal considerations

If a multi-homed SCTP end-point behind a NAT connects to a peer, it first sets up the association single-homed. Then it adds each IP address using ASCONF chunks. The address to add is the wildcard address and the lookup address also. The ASCONF chunks SHOULD also contain a vtag parameter.

10. Handling of fragmented SCTP packets

A NAT box MUST support IP reassembly of received fragmented SCTP packets. The fragments may arrive in any order.

When an SCTP packet has to be fragmented by the NAT box and the IP header forbids fragmentation a correspond ICMP packet SHOULD be sent.

11. Simplification for small NATs

Small NAT boxes, i.e. NAT boxes which only have to support a small number of concurrent SCTP associations, MAY not take the global address into account when processing packets. Therefore the global-address could also be removed from the NAT table.

This simplification may make implementing a NAT box easier, however, the collision probability is higher than using a mapping which takes the global address into account.

12. IANA Considerations

TBD

13. Security considerations

State maintenance within a NAT is always a subject of possible Denial Of Service attack. This document recommends that at a minimum a NAT run a timer on any SCTP state so that old association state can be cleaned up.

14. Acknowledgments

The authors wish to thank Qiaobing Xie, Henning Peters, Bryan Ford, David Hayes, and Jason But for their invaluable comments.

15. References

15.1. Normative References

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC 4960, September 2007.

15.2. Informative References

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

Authors' Addresses

Randall R. Stewart
The Resource Group
1700 Pennsylvania Ave NW
Suite 56
Washington, DC 20006
USA

Phone:
Email: randall.stewart@trgworld.com

Michael Tuexen
Muenster Univ. of Applied Sciences
Stegerwaldstr. 39
48565 Steinfurt
Germany

Email: tuexen@fh-muenster.de

Irene Ruengeler
Muenster Univ. of Applied Sciences
Stegerwaldstr. 39
48565 Steinfurt
Germany

Email: i.ruengeler@fh-muenster.de

Full Copyright Statement

Intellectual Property