NSIS
Internet-Draft
Expires: January 18, 2007

A. Pashalidis
H. Tschofenig
Siemens

July 17, 2006

GIST Legacy NAT Traversal
draft-pashalidis-nsis-gist-legacynats-00.txt

Status of this Memo

Copyright Notice

Abstract

This document describes an extension to the General Internet Signalling Transport (GIST) protocol that enables the protocol to traverse different types of Network Address Translator (NAT). These NATs are assumed to not support GIST, i.e. to be "legacy" NATs. The purpose of this extension is to enable GIST hosts to correctly interpret signalling messages with respect to the data traffic they refer to, in the presence of such NATs. Note that this extension does not require changes to the format of GIST messages; it merely

requires some new behaviour for non-NAT GIST nodes.

Table of Contents

1. Introduction

Network Address Translators (NATs) modify certain fields in the IP and transport layer header of the packets that traverse them. In the context of signalling as specified by the General Internet Signalling Transport (GIST) protocol [1], this behaviour may lead to the installation of state at network nodes that may be inconsistent and meaningless with respect to the data traffic that traverses these nodes.

This document describes an extension to GIST that can be used in order for GIST signalling messages to traverse GIST-unaware NATs in a way that preserves the consistency of state that is installed in the network with respect to the data flows to which the signalling messages refer. As this extension exclusively operates at the GIST layer, it is transparent to signalling applications. The document is organised as follows. The next section introduces the terminology that is used throughout this document. Section 3 provides a detailed discussion of the NAT traversal problem and highlights certain design decisions that have to be taken when addressing the problem. Section 4 lists the assumptions on which the subsequently proposed mechanisms are based. The proposed extension is described in Section 5.

2. Terminology

The terminology, abbreviations and notational conventions that are used throughout the document are as follows.

- DR: Data Receiver, same as Flow Receiver as defined in [1]
- DS: Data Sender, same as Flow Sender as defined in [1]
- GaNAT: GIST-aware NAT - a GaNAT MAY implement a number of NSLPs.
- GIST: General Internet Messaging Protocol for Signalling [1]
- NAT: Network Address Translator
- NI: NSIS Initiator; this is the GIST node (as defined in [1]) that

initiates a signalling session for a given NSLP. The NI may or may not be identical to the DS or the DR.

- NR: NSIS Responder; this is the GIST node (as defined in [1]) that acts as the last in a sequence of nodes that participate in a given signalling session. The NR may or may not be identical to the DR or the DS.
- NSIS: Next Steps in Signalling: The name of the IETF working group that specified the family of signalling protocols of which this document is also a member. The term NSIS is also used to refer to this family of signalling protocols as a whole.

- GIST-aware: Implements GIST and MAY also implement a number of NSLPs.
- GIST-unaware: GIST-unaware, does not implement any NSLP. The term is synonymous to NSIS-unaware.
- NSLP: NSIS Signalling Layer Protocol, as defined in [1]
- downstream: as defined in [1]
- upstream: as defined in [1]
- MRI: Message Routing Information, as defined in [1]
- NLI.IA: Interface Address field of the Network Layer Information object, as defined in [1]
- <- : Assignment operator. The quantity to the right of the operator is assigned to the variable to its left.
- A.B: Element B of structure A. Example: [IP header].SourceIPAddress denotes the source IP address of an IP header.
- [data item]: This notation indicates that "data item" is a single identifier of a data structure. (Square brackets do not denote optional arguments in this document.)

3. Problem Statement

According to [1], all GIST messages between two peers carry IP addresses in order to define the data flow to which the signalling refers. Moreover, certain GIST messages also carry the IP address of the sending peer, in order to enable the receiving peer to address subsequent traffic to the sender. Packets that cross an addressing boundary, say from addressing space S1 to S2, have the IP addresses in the IP header translated from space S1 to S2 by the NAT; if GIST

payloads are not translated in a consistent manner, the MRI in a GIST packet that crosses the boundary, e.g. from address space S1 to S2, refers to a flow that does not exist in S2. In fact, the flow may be invalid in S2 because at the IP address that belongs to S1 may not be routable or invalid in S2. Moreover, the IP address of the sending peer may also be not routable or invalid in the addressing space of the receiving peer. The purpose of this document is to describe an extension that enables GIST messages to be translated in a way that is consistent with the translation that NATs apply to the IP headers of the data traffic.

A NAT may be either GIST-unaware or GIST-aware. The traversal of GIST-aware NATs is described in [2] and [3]. The subject matter of this document is the traversal of GIST-unaware NATs.

A GIST-unaware NAT cannot tell data and signalling traffic apart. The installation of the NAT binding for the signalling traffic in such a NAT occurs typically independently from the installation of the NAT binding for the data traffic. Furthermore, as the NAT cannot

associate the signalling and the data traffic, it cannot indicate that an association exists between the two NAT bindings. Therefore, in the presence of such a NAT, non-NAT GIST nodes that are located on either side of the NAT have to cope with the NAT without assistance from the NAT. This would typically require initially discovering the NAT and subsequently establishing an association between between the MRI in the signalling messages and the translated IP header in the data traffic. Due to the variety of behaviours that a GIST-unaware NAT may exhibit, establishing this association is a non-trivial task.

4. Assumptions

The discussion in this document is based on the following assumptions.

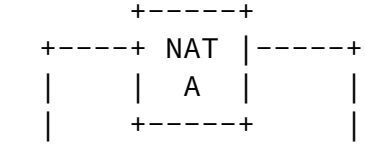
1. No IP addresses and port numbers are carried in the payloads of the NSLP. If this is not the case, then the NSLP has to provide additional mechanisms for the traversal of NATs. These mechanisms must be compatible the mechanisms described in this document.
2. The path taken by the signalling traffic between those GIST peers

that have GIST-unaware NATs in between is such that the responses to packets that a NAT sends on given interface arrive on the same interface (if such responses are sent at all).

3. The path taken by signalling traffic remains fixed between the two GIST peers, as far as the in-between NAT(s) are concerned. That is, we assume that signalling traffic traverses the same set of NATs until at least one of the following conditions is met.
   * The NSIS state that is installed at the two GIST peers expires.
   * The NSIS state that is installed at the two GIST peers is refreshed using a GIST QUERY.
   * A new GIST QUERY/RESPONSE exchange takes place due to other reasons, e.g. a detected route change.

   Note that this assumption is not necessarily met by "normal" data path coupled signalling. This is because, under "normal" data path coupled signalling, the signalling traffic is "coupled" to the data traffic at nodes that decide to act as GIST peers. Thus, under "normal" path coupled signalling, it is not always an error condition (e.g. a reason to trigger a "route change"), for example, if the set of on-path nodes, which do not act as GIST peers, changes, as long as adjacent GIST peers remain the same.
4. The data flow traverses the same set of NATs as the signalling traffic. By assumption 3, this set of NATs is fixed until the next GIST QUERY/RESPONSE procedure is executed.

5. The path-coupled routing method is used by the NSLP. (Other routing methods are not considered in this version of this document.)
6. The legacy NAT does not drop IP packets with a Router Alert Option (RAO) or an IPv6 extensions header. Furthermore, the RAO or extension header is also present in the forwarded packet. If the NAT does not do this, then there is no way for a GIST QUERY to traverse the NAT, which is a prerequisite for the mechanisms described in this document.



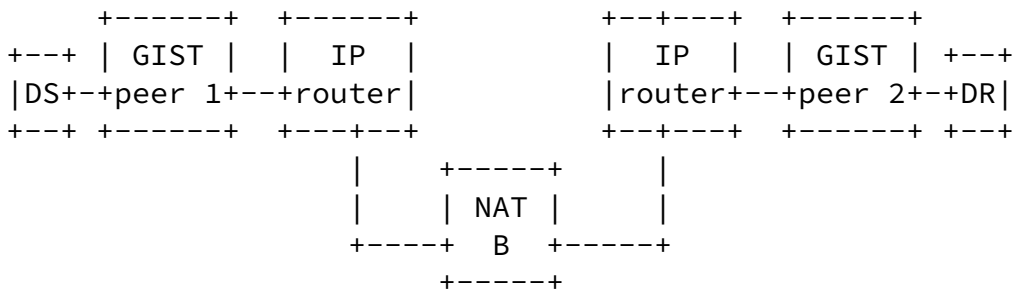


Figure 1: Network with more than one NAT at an addressing boundary

Figure 1 illustrates the importance of assumptions (3) and (4). With regard to that figure, suppose that a (D-mode) signalling session has been setup between the two adjacent GIST peers 1 and 2 and that both signalling and data traffic follows the path GIST peer 1 -> IP router -> NAT A -> IP router -> GIST peer 2. Suppose now that, after some time, GIST peer 1 decides to set up a C-mode connection with peer 2. Suppose moreover that the left IP router decides to forward the C-mode signalling traffic on the link towards NAT B. Thus, signalling traffic now follows the alternative path GIST peer 1 -> IP router -> NAT B -> IP router -> GIST peer 2. Note that this change in forwarding between the two adjacent GIST peers does not trigger a "route change" at the GIST layer because (a) it does not necessarily destroy the adjacency of peer 1 and 2 and (b) it does not necessarily destroy the coupling of the path taken by signalling traffic to that taken by data traffic (at GIST nodes). Nevertheless, assumptions (3) and (4) mandate that this situation does not occur. However, even if such a situation occurs, the mechanisms described in this document may still work as state expires after a certain timeout period.

Assumptions (2), (3) and (4) hold if, at an addressing boundary, only one NAT exists. Due to security and management reasons, this is likely to be the case in many settings.

5. Legacy NAT Traversal Mechanism

GIST-unaware NATs do not distinguish between signalling traffic and data traffic. Furthermore, the behaviour of a GIST-unaware NAT with respect to the establishment and the configuration of NAT bindings may vary significantly from one NAT to another [4]. There exists no means by which a GIST peer can predict how such a NAT will configure a future binding. A GIST-unaware NAT allocates the binding that will

be used for the data traffic independently from the binding that is used for the signalling traffic. Thereby the mapping of signalling messages to data traffic is destroyed, and cannot be re-established by GIST nodes.

The idea of enabling GIST traffic to traverse GIST-unaware NATs is somewhat similar to the mechanisms on which Teredo [6] and the STUN relay service [5], [7] are based. The idea is to tunnel signalling and data traffic over UDP, such that both data and signalling traffic use a single NAT binding. The GIST peer that is located on the other side of the NAT then removes the outer headers and also performs network address translation for both the signalling traffic (including GIST payloads) and the data traffic, in a consistent manner.

Note that two types of GIST-unaware NATs have to be dealt with, namely those that are located at the NSIS initiator (NI-side), and those that are located at the NSIS responder (NR-side). This distinction arises due to the fact that NR-side NATs are likely to drop traffic that does not match an existing binding. By contrast, NI-side NATs typically create a new binding if no matching one is found.

5.1. Traversal of NI-side legacy NATs

+----+ QUERY +----------+ QUERY +----+

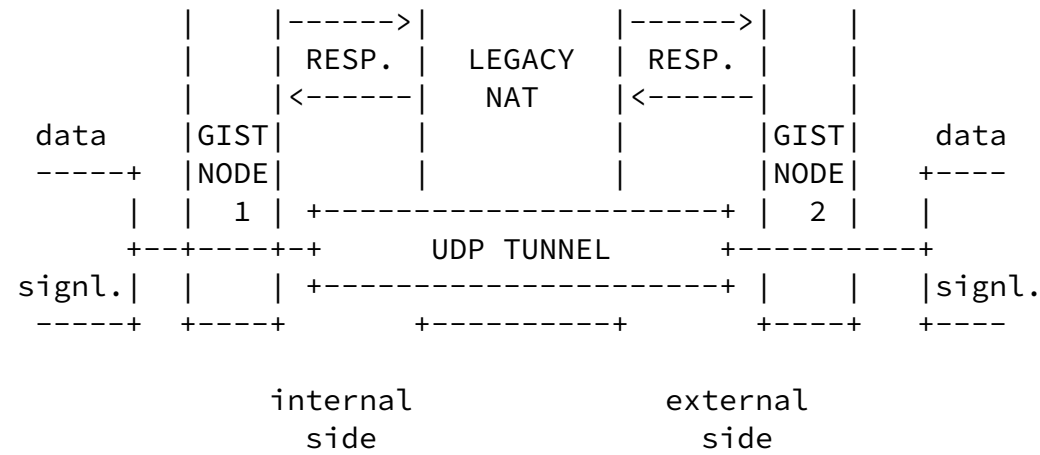


Figure 2: High level overview of NI-side legacy NAT traversal mechanism

The following may serve as indications for the existence of one or more GIST-unaware NAT(s) between two GIST peers. For the purposes of the discussion in this section, these peers are called the "upstream" GIST peer (which also happens to be the querying peer) and the "downstream" GIST peer (which also happens to be the responding peer). These indications can only be detected by the receiver of a GIST message, i.e. by the downstream peer. The first occasion these indications may be detected is with the reception of a GIST QUERY where the downstream peer assumes the role of the responder. Note that, by assumption 6, the GIST QUERY is received by the responder. Also note that != denotes inequality.

- The MRI.SourceIPAddress does not belong to the addressing space of the responding peer.
- The MRI.DestinationIPAddress does not belong to the addressing space of the responding peer.
- The IP address in the NLI.IA field does not belong to the addressing space of the responding peer.
- The S flag is set and [IP header].SourceIPAddress != NLI.IA.
- This is a GIST QUERY and [IP header].DestinationIPAddress != MRI.DestinationIPAddress.

Suppose, after detecting one or more of the above, the downstream GIST peer believes that a NAT is located between itself and the sender of the GIST QUERY. If this is indeed the case, then the NAT will have installed a UDP NAT binding as a result of the QUERY passing through it. The downstream GIST node constructs a D-mode RESPONSE according to the following rules.

- The [IP header].SourceAddress is set equal to the responder's IP address, i.e. [IP header].SourceAddress is set equal to NLI.IA.
- As a result of the above, the S flag in the RESPONSE is set.
- The MRI.[Source IP Address] field is replaced with [IP header].SourceAddress of the QUERY, i.e. the public address of the NAT.
- One stack proposal is added to the end of the set of proposals that are included in the RESPONSE by default. This last item is a proposal for UDP. It will be used for UDP tunnelling. The GIST node must be prepared to accept IP traffic that is tunneled over UDP on the advertised port.

The first of the above measures ensure that, even if the NAT exhibits an "Address and Port Dependent Filtering" behaviour, as defined in [4], the RESPONSE is not dropped by the NAT. Note that this is the most restrictive filtering behaviour, and that, therefore, the mechanism works also with less restrictive NATs.

The upstream GIST peer, on reception of the RESPONSE can also deduce that a GIST-unaware NAT is likely to be located between itself and the downstream GIST peer. This is possible because of the discrepancy of SourceAddress in the MRI sent in the QUERY and the one received in the RESPONSE. From that point onwards the upstream GIST peer tunnels both the GIST messages that belong to the current signalling session, and the data traffic to which they refer, over a UDP tunnel that it sets up with the downstream GIST peer on the advertised port. That is, the packets that the upstream GIST peer sends are such that the outer IP header is followed by a UDP header, which is in turn followed by an inner IP header. The same applies to the downstream GIST peer. In order to set up a messaging association, the upstream GIST node removes the last UDP proposal from the set of proposal received in the RESPONSE, and selects a profile as usual.

For every packet that the downstream GIST peer receives on the advertised UDP port it checks that [Outer IP header].SourceAddress is equal to [IP header].SourceAddress of the QUERY in response to which the UDP port was advertised (i.e. the public address of the NAT). If this check fails, the packet is silently discarded. Note that, in order to perform this check, the peer needs to allocate some state before a CONFIRM message is received. This state, however, is not necessarily per-session state, and various DoS exposure mitigation techniques can be applied if the peer finds itself heavily loaded. One such measure is to temporarily turn off the support of GIST-unaware NAT traversal.

A packet that the downstream peer receives over the tunnel is either a GIST message or a data packet. It is a GIST message if [inner IP

header].DestinationAddress belongs to the peer (i.e. is equal to the one advertised in the NLI.IA field), and a data packet otherwise. Note that, if the downstream peer is also the DR, then this distinction does not apply. However, in this case the inner transport layer header can be used to unambiguously determine whether the packet belongs to an established GIST messaging association, or a data flow.

5.1.1. Treatment of Data Traffic

When the downstream GIST peer receives a data packet P from the upstream GIST peer over the UDP tunnel, it should ensure that the following conditions are met.

- The inner [IP header].[transport protocol] field and the inner [Transport layer header].DestinationPort of P matches the MRI of a GIST QUERY in response to which the UDP tunnel parameters were advertised. Note that multiple such GIST queries may exist if the same tunnel is used for multiple signalling sessions (and therefore multiple data flows) between the upstream and the downstream GIST peers.
- If the signalling session is to run over a secure connection (e.g. IPsec, TLS), and if required by local policy, then the messaging association has been established. That is, local policy may dictate that P MUST NOT be forwarded until the messaging association establishment has been completed sucessfully.

The downstream GIST peer then removes the outer IP and UDP headers, replaces [inner IP header].SourceAddress with an IP address denoted IPTRANS. This IP address must be chosen in a way that ensures that packets sent to IPTRANS will arrive at the downstream GIST peer. IPTRANS MUST thus be one of the GIST peer's own IP addresses (preferably, but not necessarily, bound to the interface over which P will be forwarded), unless in an exceptional situation, explained below. The downstream GIST peer MAY also replace [inner transport header].SourcePort with a different source port, denoted PORTTRANS. The resulting data packet is forwarded according to the peer's routing table.

A data packet K that arrives from the downstream direction, which belongs to same bidirectional data flow as P but flows in the opposite direction (i.e. from DR to DS), MUST be mapped to the correct UDP tunnel towards the upstream GIST peer. Moreover, the upstream peer (which is the tunnel endpoint) MUST be able to demultiplex multiple data flows that may arrive over the same tunnel. To this end, the downstream GIST peer MUST use a unique (IPTRANS, PORTTRANS) pair for each tunelled data flow. Note that, in the situation where the number of NATs over which the downstream GIST

node entertains tunnels, exceeds the number of IP addresses that the downstream peer may chose IPTRANS from, then the number of tunnels may exceed the number of available (IPTRANS,PORTTRANS) pairs (for a given transport layer protocol). The same may happen in the presence of tunnels over which multiple data flows are tunnelled. In such an exceptional situation, i.e. when the downstream GIST runs out of (IPTRANS, PORTTRANS) pairs (for a given transport layer protocol), then it MAY use the public address of the NAT, behind which the data flow originates, as IPTRANS. It should be noted that, in this case, routing assymmetry on the path that the packet K takes on its way to the downstream GIST node may cause the mechanism to fail, because K may never arrive at the downstream GIST node.

When the downstream GIST peer receives a data packet K, it looks at K.[IP header].DestinationAddress, K.[IP header].Protocol, and K.[Transport layer header].DestinationPort. If the downstream GIST node behaves as explained above, this triple uniquely defines the tunnel, even in the presence of multiple NATs (possibly from multiple domains), and multiple tunnels per NAT, as explained above. Before the downstream GIST node forwards K over the tunnel to the upstream GIST node, it MUST translate K.[IP header].DestinationAddress and, if applicable, K.[IP header].DestinationPort according to the translation applied to P. If the aforementioned triple does not match an existing tunnel, then normal processing applies to K (whatever that means at the downstream GIST node).

Finally, note that, if the data flow exists before signalling is initiated, then the application may be adversely affected by the mechanism described in this section. This is because, prior to signalling, the DR sees the public address of the NAT as the address of the DS. However, subsequent to signalling (and the associated tunnelling), the DR will see IPTRANS as the address of the DS (and

may therefore assume that the DS has changed). In order to avoid this, the downstream GIST peer could use the public address of the NAT as IPTRANS if the data traffic already flows at the time that signalling is initiated. In order, however, to select the correct value for PORTTRANS, the downstream GIST peer must be able to correlate the data traffic before tunneling and after tunnelling. The source port in the inner transport layer header of a tunneled data packet X that belongs to a given flow, is equal to the source port of an non-tunneled data packet X' that belongs to the same flow, if and only if the NAT binding over which X' travels is source port preserving. Unfortunately, legacy NATs do not always install such bindings [4]. It is NOT RECOMMENDED for the downstream GIST peer to assume that NAT bindings are source port preserving. Therefore, the mechanism described in this section assumes that data traffic flows after signalling state has been setup in the network.

5.1.2. Treatment of Signalling Traffic

The processing of GIST messages that arrive over a UDP tunnel adheres to the usual rules once the outer IP and UDP headers are stripped off. For GIST messages that the downstream GIST peer sends towards the upstream direction, the correct IP/UDP encapsulation must be used. To this end, the peer must keep the necessary state in association with the routing state for the upstream GIST peer. For GIST messages that are sent towards the downstream direction, the GIST peer must also change the MRI such that it reflects the translation for the data traffic. That is, the peer MUST set MRI.SourceIPAddress = IPTRANS and, if applicable, MRI.SourcePort=PORTTRANS.

5.1.3. Refreshing NSIS State

According to [1], NSIS signalling state must be refreshed regularly. To this end, the NI periodically sends GIST QUERY messages which are forwarded along the data path. When the upstream GIST peer receives such a QUERY (i.e. a QUERY that has the purpose of refreshing signalling state), and if a UDP tunnel already exists for the data traffic to which this QUERY refers to, then the upstream peer MUST send this QUERY as if no tunnel existed, i.e. as described above. This is in order to enable a new GIST node to be identified as the downstream peer. However, the upstream GIST peer SHOULD use the same

source port in the refreshing QUERY as in the outer UDP header of the tunnelled packets. This is in order to turn the NSIS refresh mechanism into a mechanism that keeps the UDP NAT binding alive. This is important if no data traffic is sent for an extended period of time.

If the downstream GIST peer receives a GIST QUERY for which a tunnel and a GIST messaging association already exists, then it send the response over the existing messaging association, in accordance with [1]. This involves tunnelling the RESPONSE over the tunnel, as descibed above. If the downstream GIST peer receives a GIST QUERY for which a tunnel, but no messaging association exists yet, then, if policy permits, the peer sends the RESPONSE as if no tunnel existed. In this RESPONSE, the peer MAY propose the same tunnel parameters as in the original RESPONSE.

5.2. Traversal of NR-side legacy NATs

The traversal of NR-side legacy NATs is not as straight-forward as the case of NI-side legacy NAT traversal. This is because an NR-side legacy NAT is likely to block all "unsolicited" incoming traffic. That is, such a NAT is unlikely to install a NAT binding on the basis of a packet that arrives on its public side. Instead, such packets

are typically only forwarded towards the private side, if they match an already installed NAT binding.

The NR-side legacy NAT traversal mechanism will be specified in a future version of this document.

6. Security Considerations

Editor's note: this section will be completed after normative behaviour has been fully specified.

7. Acknowledgements

The authors would like to thank Robert Hancock and Martin Stiemerling for his insightful comments.

8. References

8.1. Normative References

[1] Schulzrinne, H. and R. Hancock, "GIST: General Internet Signalling Transport", draft-ietf-nsis-ntlp-09 (work in progress), February 2006.

[2] Stiemerling, M., Tschofenig, H., and C. Aoun, "NAT/Firewall NSIS Signaling Layer Protocol (NSLP)", draft-ietf-nsis-nslp-natfw-09 (work in progress), February 2006.

[3] Pashalidis, A. and H. Tschofenig, "GIST NAT Traversal", draft-pashalidis-nsis-gimps-nattraversal-03.txt (work in progress), June 2006.

[4] Audet, F. and C. Jennings, "NAT Behavioral Requirements for Unicast UDP", draft-ietf-behave-nat-udp-07 (work in progress), May 2006.

[5] Rosenberg, J., Tschofenig, H., Huitema, C., Mahy, R., and D. Wing, "Simple Traversal of UDP Through Network Address Translators (NAT) (STUN))", draft-ietf-behave-rfc3489bis-03 (work in progress), February 2006.

8.2. Informative References

[6] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, February 2006.

[7] Rosenberg, J., Mahy, R., and C. Huitema, "Obtaining Relay Addresses from Simple Traversal of UDP Through NAT", draft-ietf-behave-turn-01.txt (work in progress), February 2006.

Authors' Addresses

Andreas Pashalidis
Siemens

Otto-Hahn-Ring 6
Munich, Bavaria 81739
Germany

Email: Andreas.Pashalidis@siemens.com

Hannes Tschofenig
Siemens
Otto-Hahn-Ring 6
Munich, Bavaria 81739
Germany

Email: Hannes.Tschofenig@siemens.com

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.