NSIS Working Group
Internet-Draft
Expires: November 19, 2004

M. Stiemerling
NEC
H. Tschofenig
Siemens
M. Martin
NEC
C. Aoun
Nortel Networks

May 21, 2004

NAT/Firewall NSIS Signaling Layer Protocol (NSLP)
draft-ietf-nsis-nslp-natfw-02

Status of this Memo

Copyright Notice

Abstract

This memo defines the NSIS Signaling Layer Protocol (NSLP) for Network Address Translators and Firewalls. This NSLP allows hosts to signal along a data path for Network Address Translators and Firewalls to be configured according to the data flow needs. The network scenarios, problems and solutions for path-coupled Network Address Translator and Firewall signaling are described. The overall

architecture is given by the framework and requirements defined by Next Steps in Signaling (NSIS) working group. This is one of two NSIS Signaling Layer Protocols (NSLPs) the working group will address during its work.

Table of Contents

1. Introduction

Firewalls and Network Address Translators (NAT) have been both used throughout the Internet for many years and they will be present in future. Using Firewalls brings security to networks and in times of IPv4 address depletion NATs virtually extend IP address space. In general, both types may be obstacles to many applications, since they only allow specific applications to traverse them (i.e., HTTP traffic or in general client/server applications). Other applications, for instance, IP telephony or any other peer-to-peer application, with more dynamic properties suffer from Firewalls and NATs so that they do not work at all. Therefore, many applications cannot traverse Firewall or NATs.

Several solutions to enable any application to traverse those boxes have been proposed and are currently used. Typically, application level gateways (ALG) have been integrated and so configuring Firewalls and NATs dynamically. Another approach is middlebox communication (MIDCOM, currently under standardization at the IETF). In this approach Firewall and NAT external ALGs configure them via the MIDCOM protocol [7]. Several other work around solutions are available as well, see STUN [32] and [31]. However, all of these approaches introduce other problems that are hard to solve; like dependencies on certain NAT implementations or dependency on topology.

NAT and Firewall (NATFW) signaling share a property with Quality of Service (QoS) signaling, i.e., in both cases it is required to reach

any device on the data path that is involved in QoS or NATFW treatment of data packets. For both, NATFW and QoS, signaling travels path-coupled, meaning that the signaling messages follow exactly the same path as the data packets do. RSVP [14] is an example for a QoS signaling protocol.

This memo defines a path-coupled signaling protocol in the framework of NSIS for NAT and Firewall configuration, called the NATFW NSIS Signaling Layer Protocol (NSLP). The general framework of NSIS is outlined in [1] and introduces the split between NSIS transport layer and NSIS signaling layer. The transport of NSLP messages is handled by NSIS Network Transport Layer Protocol (NTLP, see [3]) and takes care about NSLP message transport. The signaling logic for QoS and NATFW signaling is implemented in the different NSLPs. The QoS NSLP is defined in [4], furthermore the general requirements for NSIS are defined in [2].

There is a series of related documents to NATFW NSLP discussing several other aspects of path-coupled NATFW signaling, including security [20], migration [17], intrarealm signaling [18], and

inter-working with SIP [19].

The NATFW NSLP allows requesting the configuration of NATs and/or Firewalls along the data path to enable data flows to traverse these devices without being obstructed. A simplified example: A source host sends a NATFW NSLP signaling message towards its data destination. This message follows the data path and every NATFW NSLP NAT/Firewall along the data path intercepts these messages, processes it and configures itself accordingly. Afterwards, the actual data flow can traverse every configured Firewall/NAT.

NATFW NSLP runs in two different modes, one is the path directed mode where Firewalls and NATs are configured along the data path as pointed out in the above example. The second one is the reserve mode, where NATs are detected by the NSLP/NTLP within the network and a public reachable IP address and port number are reserved. This reserve mode enables hosts located behind NATs to receive data originated in the public Internet on the reverse data path. Both modes create NATFW NSLP and NTLP state in the network. The NSLP state is maintained via a soft-state mechanism. State includes not only signaling state, but as well as NAT bindings and Firewall rules.

This state is maintained via a lifetime and must be kept alive via a lifetime extension mechanism if needed. Two signaling messages are used for deleting state explicitly and extending state's lifetime. In general, all NATFW NSLP signaling messages are exchanged end-to-end.

Traversal of non NATFW NSLPs or the NTLP is out of scope of this document. Furthermore, only Firewalls and NATs are considered in this document, any other device, for instance IPSec security gateway, is out of scope.

Section 2 describes the network environment for NATFW NSLP signaling and highlights the required trust relationship/ authorization. Section 3 defines the NATFW signaling protocol with its message components, message formats, and protocol operations. The remaining document refers in Section 4 to transition issues and security considerations are handled in [20]. Currently unsolved problems and challenges are listed and discussed in Appendix A. Please note that readers familiar with possible locations of Firewalls and NAT in networks can safely skip Section 2.

1.1 Terminology and Abbreviations

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 RFC 2119.

This document uses terms defined in [2]. Furthermore, these following terms are used:

- NSIS NAT Forwarding State: The term "NSIS NAT Forwarding State" in this context refers to a state used to forward the NSIS signaling message beyond the targeted destination address; that state is typically used when the NSIS Responder address is not known
- Sender-/Receiver Initiated Signaling
  Sender-initiated: NAT bindings and Firewall rules are created immediately when the "path" message hits the NSIS nodes. With "path" message we refer to the signaling message traveling from the data sender towards the data receiver.
  Receiver-initiated: NAT bindings and Firewall rules are created when the "reserve" message returns from the other end. With "reserve" message we refer to a signaling message on the

reverse path, this means from the receiver to the sender (i.e. backwards routed).
Note that these definitions have nothing to do with number of roundtrips, who performs authorization etc.

- Policy rule: In general, a policy rule is "a basic building block of a policy-based system. It is the binding of a set of actions to a set of conditions - where the conditions are evaluated to determine whether the actions are performed." [RFC3198]. In the context of NSIS NATFW NSLP the condition is a specification of a set of packets to which rules are applied. The set of actions always contains just a single element per rule, and is limited to either action "reserved" or action "enable".
- Firewall: A packet filtering device that matches packet against a set of policy rules and applies the actions. In the context of NSIS NATFW NSLP we refer to this device as Firewall.
- Network Address Translator: Network Address Translation is a method by which IP addresses are mapped from one realm to another, in an attempt to provide transparent routing to hosts (see [9]). Network Address Translators are devices that perform this method.
- Middlebox: from [12]: "A middlebox is defined as any intermediate device performing functions other than the normal, standard functions of an IP router on the datagram path between a source host and a destination host". The term middlebox in context of this document and in NSIS refers to Firewalls and NATs only. Other types of middlebox are currently outside the scope.
- Security Gateway: IPsec based gateways.
- NSIS Initiator (NI): the signaling entity, which makes the resource request, usually as a result of user application request.
- NSIS Responder (NR): the signaling entity , which acts as the final destination for the signaling and can optionally interact with applications as well.
- NSIS Forwarder (NF): the signaling entity between an NI and NR which propagates NSIS signaling further through the network.

- Receiver (DR or R): the node in the network, which is receiving the data packets of a flow.
- Sender (DS or S): the node in the network, which is sending the data packets of a flow.
- NATFW NSLP session: Application layer flow of information for which some network control state information is to be manipulated or monitored (as defined in [1]). The control state for NATFW

NSLP is NSLP state and associated policy rules at the middlebox.

- NSIS peer or peer: NSIS node with which a NSIS adjacency has been created as defined in [3].
- Edge NAT: By edge NAT we refer to the NAT device, which is reachable from outside and has a globally routable IP address.
- Public Network: Definition according to [8] is "A Global or Public Network is an address realm with unique network addresses assigned by Internet Assigned Numbers Authority (IANA) or an equivalent address registry. This network is also referred as External network during NAT discussions."
- Private/Local Network: Definition according to [8] is " A private network is an address realm independent of external network addresses. Private network may also be referred alternately as Local Network. Transparent routing between hosts in private realm and external realm is facilitated by a NAT router." IP address space allocation for private networks is recommended in [33]
- Public/Global IP address: An IP address located in the public network.
- Private/Local IP address: An IP address located in the private network.

1.2 Middleboxes

The term middlebox raises different expectations about functionality provided by such a device. Middleboxes in the scope of this memo are Firewalls that filter data packets against their set of filter rules and NATs that translate addresses from one address realm to another address realm. Other types of middleboxes, for instance QoS traffic shapers and security gateways, are out of scope.

The term NAT used in this document is placeholder for a range of different NAT flavors. We consider those types of NATs:

- traditional NAT (basic NAT and NAPT)
- Bi-directional NAT
- Twice-NAT
- Multihomed NAT

For a detailed discussion about each NAT type please see [8].

Both types of middleboxes use policy rules for decision on data packet treatment. Policy rules consist of a 5-tuple and an associated action. Data packets matching this 5-tuple experience the

policy rule action. A 5-tuple consists of:
- Source IP address and port number
- Destination IP address and port number
- Transport protocol

Actions for Firewalls are usually:
- Allow: forward data packet
- Deny: block data packet and discard it
- Other actions like logging, diverting, etc

Actions for NATs are (amongst many others):
- Change source IP address and port number to a global routeable IP address and port number.
- Change destination IP address and port number to a private IP address and port number.

The exact implementation of policy rules and mapping to Firewall rule sets and NAT bindings or sessions at the middlebox is an implementation issue and thus out of scope of this document.

Some devices entitled as Firewalls only accept traffic after cryptographic verification (i.e. IPsec protected data traffic). Particularly for network access scenarios either link layer or network layer data protection is common. Hence we do not address these types of devices (referred as security gateways) since per-flow signaling is rather uncommon in this environment. For a discussion of network access authentication and associated scenarios the reader is referred to the PANA working group (see [26]).

Discovering security gateways, which was also mentioned as an application for NSIS signaling, for the purpose of executing an IKE to create an IPsec SA, is already solved without requiring NSIS.

In mobility scenarios an often experienced problem is the traversal of a security gateway at the edge of the corporate network. Network administrators often rely on the policy that only authenticated data traffic is allowed to enter the network. A problem statement for the traversal of these security gateways in the context of Mobile IP can be found at [25]).

Other proposals for path-coupled NAT and Firewall traversal like RSVP and CASP are described in [27] and [28].

1.3 General Scenario for NATFW Traversal

The purpose of NSIS NATFW signaling is to enable any communication between endpoints across networks even in presence of middleboxes. It is expected that those middleboxes be configured in such a way that NSIS NATFW signaling messages itself are allowed to traverse them. NSIS NATFW NSLP signaling is used to install such policy rules

in all middleboxes along the data path. Firewalls are configured to forward data packets matching the policy rule provided by the NSLP signaling. NATs are configured to translate data packets matching the policy rule provided by the NSLP signaling.

The basic high-level picture of NSIS usage is that endhosts are located behind middleboxes (NAT/FW in Figure 1). Applications located at these endhosts try to establish communication between them and use NSIS NATFW NSLP signaling to establish policy rules on a data path, which allows the said data to travel from the sender to the receiver unobstructed. The applications can somehow trigger middlebox traversal (e.g. via an API call) at the NSIS entity at the local host.

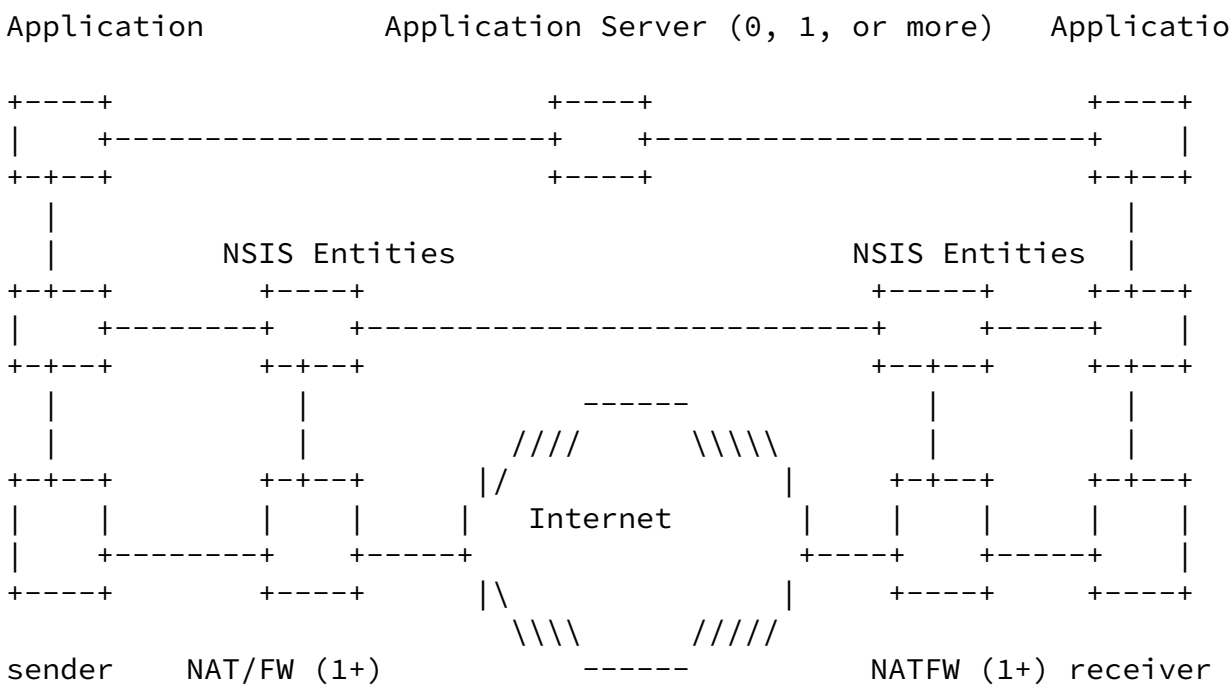


Figure 1: Generic View on NSIS in a NAT / Firewall case

For running NATFW signaling it is necessary that each Firewall and each NAT involved in the signaling communication runs an NSIS NATFW entity. There might be several NATs and FWs in various possible combinations on a path between two hosts. The reader is referred to Section 2.1 where different scenarios are presented.

2. Network Environment

2.1 Network Scenarios for Protocol Functionality

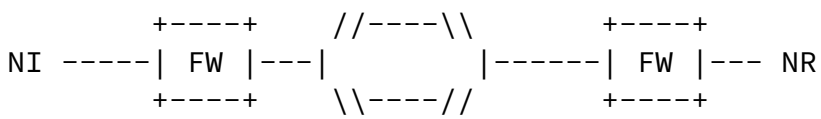
This section introduces several scenarios for middleboxes in the Internet. Middleboxes are located at different locations, i.e. at Enterprise network borders, within enterprise networks, mobile phone network gateways, etc. In general, middleboxes are placed more towards the edge of networks and less in network cores. Those middleboxes are not only either Firewall or NAT and any other type of combination is possible. Thus, combined Firewall and NATs are available.

NSIS initiators (NI) are sending NSIS NATFW NSLP signaling messages via the regular data path to the NSIS responder (NR). On the data path NATFW NSLP signaling messages reach different NSIS peers that have the NATFW NSLP implemented. Each NATFW NSLP node processes the signaling messages according to Section 3 and installs, if necessary, policy rules for subsequent data packets.

Each following section introduces a different scenario for a different set of middleboxes and their ordering within the topology. It is assumed that each middlebox implements the NSIS NATFW NSLP signaling protocol.

2.1.1 Firewall traversal

This section describes a scenario with Firewalls only and NATs are not involved. Both end hosts are behind a Firewall that is connected via the public Internet. Figure 2 shows the topology. The part labeled "public" is the Internet connection both Firewalls.



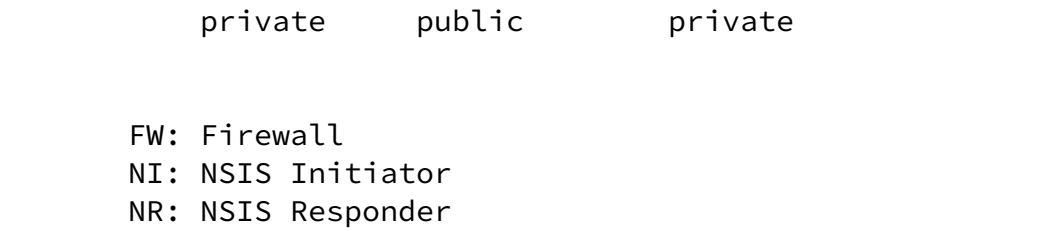


Figure 2: Firewall Traversal Scenario

Each Firewall on-path must provide traversal service for NATFW NSLP

in order to permit the NSIS message to reach the other end host. All Firewalls process NSIS signaling and establish appropriate policy rules, so that the required data packet flow can traverse them.

2.1.2 NAT with two private Networks

Figure 3 shows a scenario with NATs at both ends of the network. Therefore, each application instance, NSIS initiator and NSIS responder, are behind NATs. The outermost NAT at each side is connected to the public Internet. The NATs are labeled as MB (for middlebox), since those devices implement at least NAT-only, but can implement Firewalling as well.

Only two middleboxes MB are shown in Figure 3 at each side, but in general more than one MB on each side must be considered.

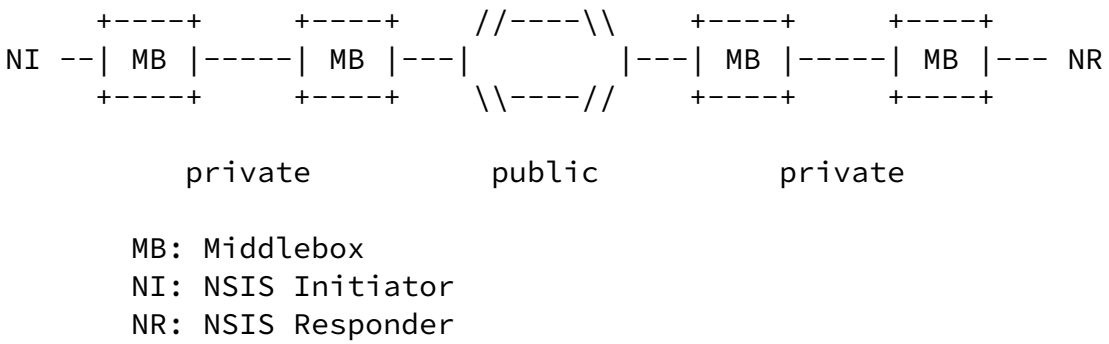


Figure 3: NAT with two private networks Scenario

Signaling traffic from NI to NR has to traverse all four middleboxes on the path and all four middleboxes must be configured properly to allow NSIS signaling to traverse. The NATFW signaling must configure all middleboxes and consider any address translation in further signaling. The sender (NI) has to know the IP address of the receiver (NR) in advance, otherwise he cannot send a single NSIS signaling message towards the responder. Note that this IP address is not the private IP address of the responder. Instead a NAT binding (including a public IP address) has to be obtained from the NAT that subsequently allows packets hitting the NAT to be forwarded to the receiver within the private address realm. This generally requires further support from an application layer protocol for the purpose of discovering and exchanging information. The receiver might have a number of ways to learn its public IP address and port number and might need to signal this information to the sender using the application level signaling protocol.

2.1.3 NAT with private network on sender side

This scenario shows an application instance at the sending node that is behind one or more NATs (shown as MB). The receiver is located in the public Internet.

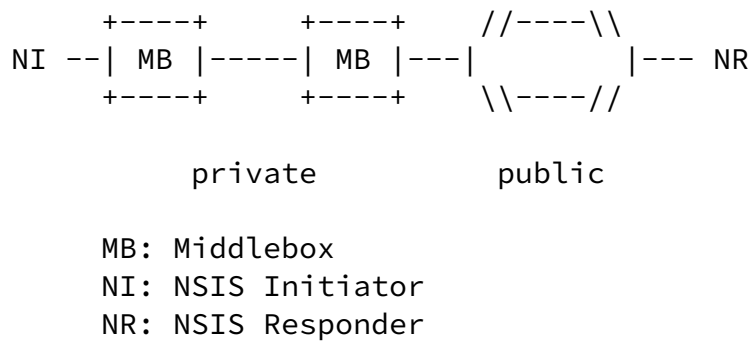


Figure 4: NAT with private network on sender scenario

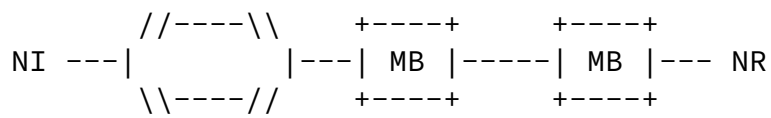
The traffic from NI to NR has to traverse only middleboxes on the sender's side. The receiver has a public IP address. The NI sends

its signaling message directly to the address of the NSIS responder. Middleboxes along the path intercept the signaling messages and configure the policy rules accordingly.

Note that the data sender does not necessarily know whether the receiver is behind a NAT or not, hence, it is the receiving side that has to detect whether itself is behind a NAT or not. As described in Section 3.2.2 NSIS can also provide help for this procedure.

2.1.4 NAT with private network on receiver side

The application instance receiving data is behind one or more NATs.



MB: Middlebox
NI: NSIS Initiator
NR: NSIS Responder

Figure 5: NAT with private network on receiver Scenario

Initially, the NSIS responder must determine its public reachable IP

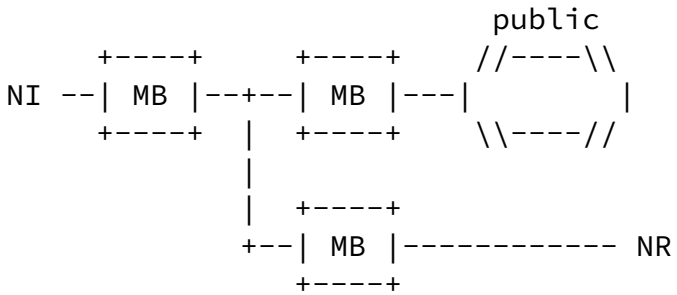
address at the external middlebox and notify the NSIS initiator about this address. One possibility is that an application level protocol is used, meaning that the public IP address is signaled via this protocol to the NI. Afterwards the NI can start its signaling towards the NR and so establishing the path via the both middleboxes MB.

This scenario describes the use case for the reserve mode of the NATFW NSLP.

2.1.5 Both End Hosts behind twice-NATs

This is a special case, where the main problem is to detect that both nodes are logically within the same address space, also behind a twice-NAT (see [8] for discussion about twice-NAT functionality).

Sender and receiver are both within a private address realm and potentially have overlapping IP addresses. Figure 6 shows the ordering of NATs. This is a common configuration in several networks, particularly after the merging of companies that have used the same address space, thus having overlapping addresses in many cases.



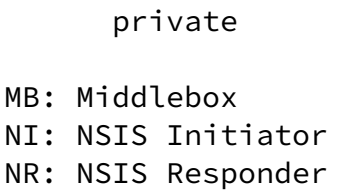


Figure 6: NAT to public, sender and receiver behind twice-NAT Scenario

The middleboxes shown in Figure 6 are twice-NATs, i.e. they map IP addresses and port numbers on both sides, at private and public interfaces.

This scenario requires assistance of application level entities, like DNS server. Those application level gateways must handle request that are based on symbolic names and configure the middleboxes so that data packets are correctly forwarded from NI to NR. The configuration of those middleboxes may require other middlebox communication protocols, like MIDCOM [7]. NSIS signaling is not required in the twice-NAT only case, since the middleboxes of type twice-NAT are configured by other means. Nevertheless, NSIS signaling might by useful when there are Firewalls on path. In this case NSIS will not configure any policy rule at twice-NATs, but will configure policy rules at the intermediate Firewalls. The NSIS signaling protocol must be at least robust enough to survive this scenario.

2.1.6 Both End Hosts behind same NAT

When NSIS initiator and NSIS responder are behind the same NAT (thus being in the same address realm, see Figure 7), they are most likely not aware of this fact. As in Section 2.1.4 the NSIS responder must determine its public IP address in advance and transfer it to the NSIS initiator. Afterwards, the NSIS initiator can start sending the signaling messages to the responder's public IP address. During this process, a public IP address will be allocated for the NSIS initiator at the same middlebox as for the responder. Now, the NSIS signaling and the subsequent data packets will traverse the NAT two times: from

initiator to public IP address of responder (first time) and from

public IP address of responder to responder (second time). This is the worst case, both sender and receiver obtain a public IP address at the NAT and the communication path is not optimal anymore.

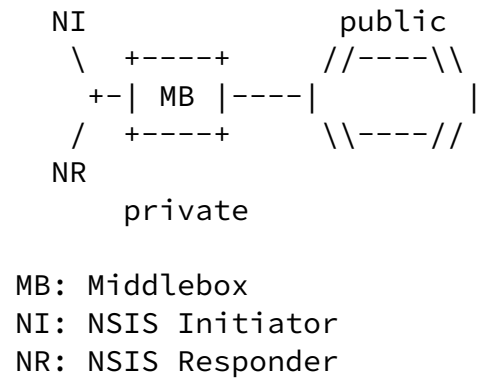


Figure 7: NAT to public, both host behind same NAT

NSIS NATFW signaling protocol should support mechanisms to detect such a scenario. The signaling should directly by exchanged between NI and NR without involving the middlebox.

2.1.7 IPv4/v6 NAT with two private networks

This scenario combines the usage case mentioned in Section 2.1.2 with the IPv4 to IPv6 transition scenario, i.e. using Network Address and Protocol Translators (NAT-PT, [11]).

The difference to the other scenarios is the use of IPv6 to IPv4 (and vice versa) address and protocol translation. Additionally, the base NTLP must take care of this case for its own functionality of forwarding messages between NSIS peers.

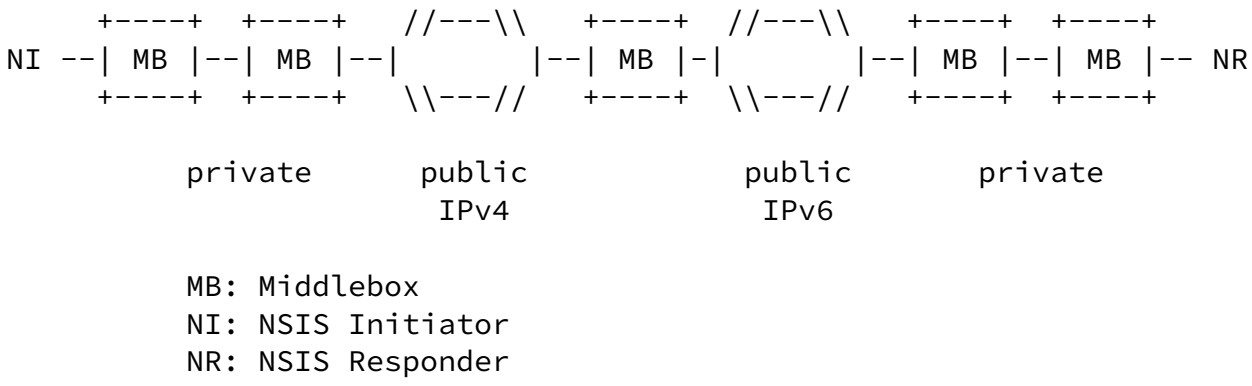


Figure 8: IPv4/v6 NAT with two private networks

This scenario needs the same type of application level support as described in Section 2.1.5 and so those issues of twice-NATs apply here as well.

2.1.8 Multihomed Network with NAT

The previous chapters sketched network topologies where NAT and Firewalls are ordered sequentially on the path. This chapter describes a multihomed scenario with two NATs to the Internet.

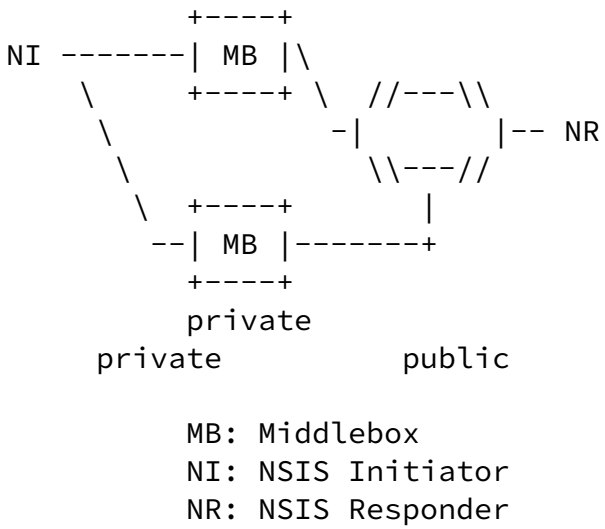


Figure 9: Multihomed Network with two NATs

Depending on the destination the one or the other middlebox is used for the data flow. Which middlebox is used depends on local routing

decisions. NATFW NSLP must be able to handle this situation proper, see Section 3.2.2 for a more elaborated discussion of this topic with

respect to NATs.

2.2 Trust Relationship and Authorization

Trust relationships and authorization are very important for the protocol machinery. Trust and authorization are closely related to each other in the sense that a certain degree of trust is required to authorize a particular action. For any action (e.g. "create/delete /prolong policy rules" then authorization is very important due to the nature of middleboxes.

It is particularly not surprising that different degrees of required authorization in a QoS signaling environment and middlebox signaling exist. As elaborated in [23], establishment of a financial relationship is very important for QoS signaling, whereas for middlebox signaling is not directly of interest. For middlebox signaling a stronger or weaker degree of authorization might be needed.

Different trust relationships that appear in middlebox signaling environments are described in the subsequent sections. Peer-to-peer trust relationships are those, which are used in QoS signaling today and seem to be the simplest. However, there are reasons to believe that this is not the only type of trust relationship found in today's networks.

2.2.1 Peer-to-Peer Trust Relationship

Starting with the simplest scenario it is assumed that neighboring nodes trust each other. The required security association to authenticate and to protect a signaling message is either available (manual configuration) or dynamically established with the help of an authentication and key exchange protocol. If nodes are located closely together it is assumed that security association establishment is easier than establishing it between far distant node. It is, however, difficult to describe this relationship generally due to the different usage scenarios and environments. Authorization heavily depends on the participating entities but for this scenario it is assumed that neighboring entities trust each

other (at least for the purpose of policy rule creation, maintenance and deletion). Note that Figure 10 does not illustrate the trust relationship between the end host and the access network.

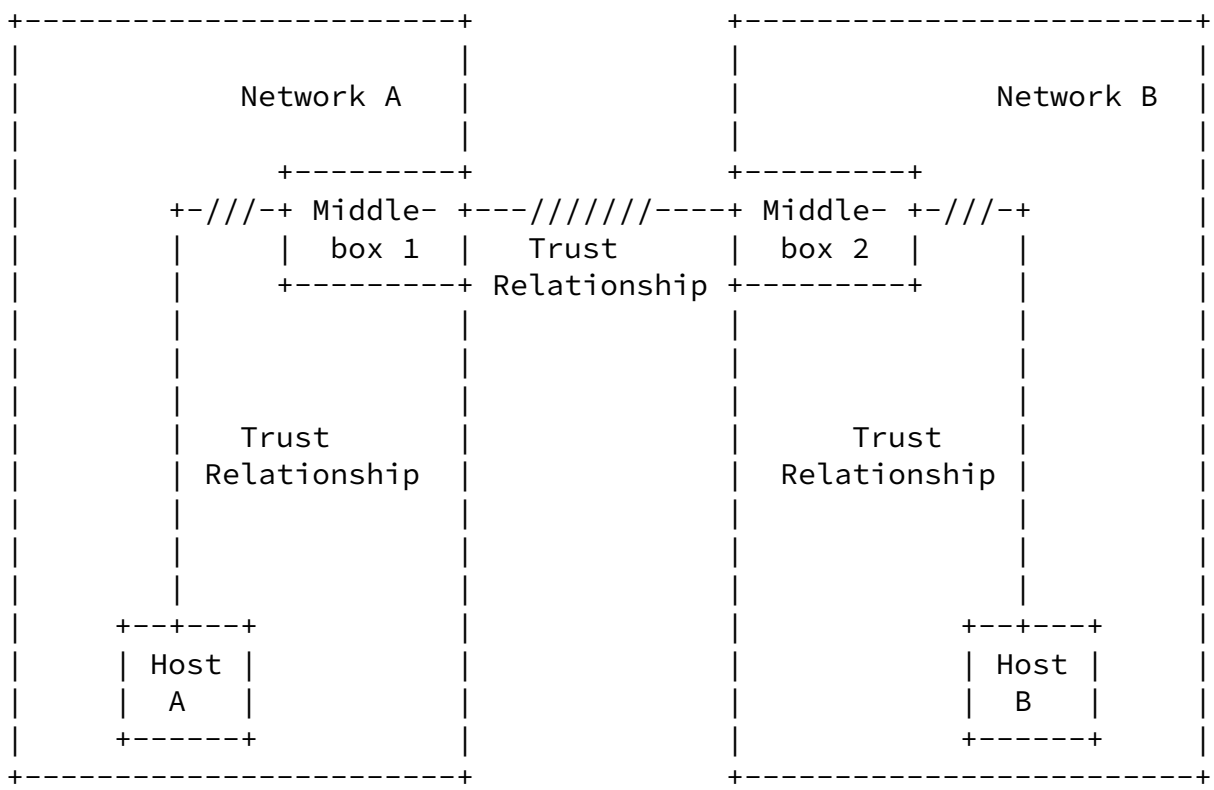


Figure 10: Peer-to-Peer Trust Relationship

2.2.2 Intra-Domain Trust Relationship

In larger corporations often more than one middlebox is used to protect different departments. In many cases the entire enterprise is controlled by a security department, which gives instructions to the department administrators. In such a scenario a peer-to-peer

trust-relationship might be prevalent. Sometimes it might be necessary to preserve authentication and authorization information within the network. As a possible solution a centralized approach could be used whereby an interaction between the individual middleboxes and a central entity (for example a policy decision point - PDP) takes place. As an alternative individual middleboxes could exchange the authorization decision to another middlebox within the same trust domain. Individual middleboxes within an administrative domain should exploit their trust relationship instead of requesting authentication and authorization of the signaling initiator again and again. Thereby complex protocol interaction is avoided. This provides both a performance improvement without a security disadvantage since a single administrative domain can be seen as a single entity. Figure 11 illustrates a network structure, which uses a centralized entity.

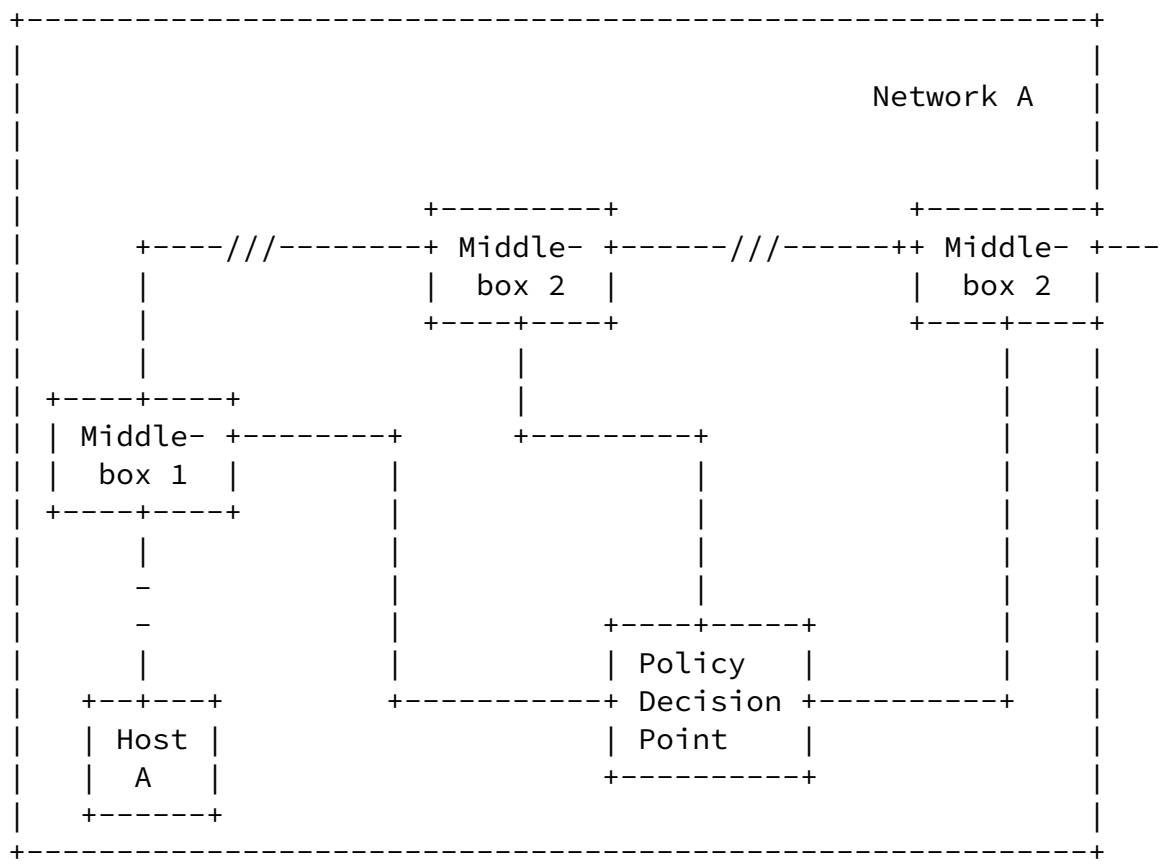
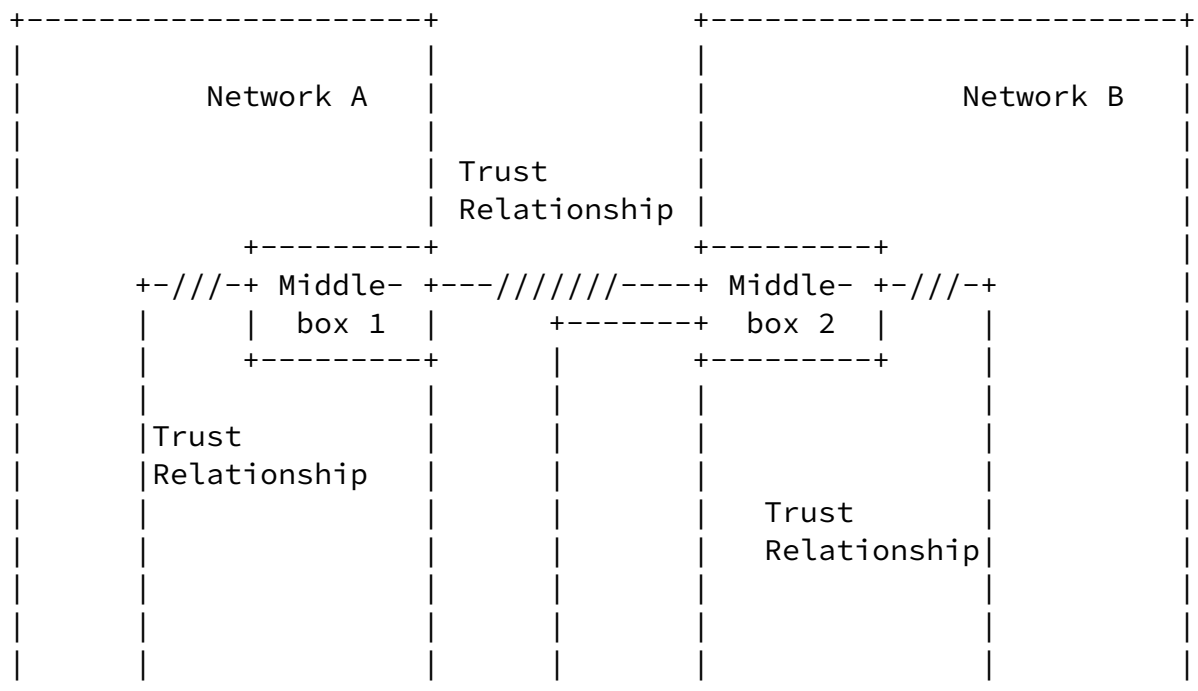


Figure 11: Intra-domain Trust Relationship

2.2.3 End-to-Middle Trust Relationship

In some scenarios a simple peer-to-peer trust relationship between participating nodes is not sufficient. Network B might require additional authorization of the signaling message initiator. If authentication and authorization information is not attached to the initial signaling message then the signaling message arriving at Middlebox 2 would cause an error message to be created, which indicates the additional authorization requirement. In many cases the signaling message initiator is already aware of the additionally required authorization before the signaling message exchange is executed. Replay protection is a requirement for authentication to the non-neighboring middlebox, which might be difficult to accomplish without adding additional roundtrips to the signaling protocol (e.g. by adding a challenge/response type of message exchange).

Figure 12 shows the slightly more complex trust relationships in this scenario.



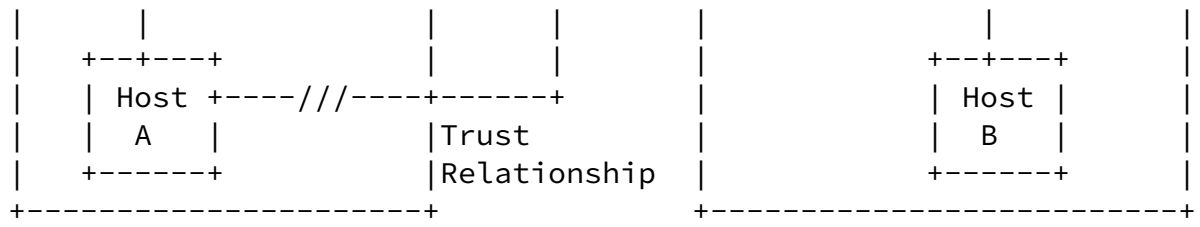


Figure 12: End-to-Middle Trust Relationship

3. Protocol Description

The protocol description section defines the NSIS NATFW NSLP with its messages, objects, and the protocol semantics. Section 3.1 introduces the protocol and Section 3.3 defines the syntax of the messages and objects. The protocol behavior is defined in Section 3.2.

3.1 Basic protocol overview

The NSIS Signaling Layer Protocol (NSLP) for NAT and FW traversal is

carried over the NSIS Transport Layer Protocol (NTLP) defined in [3]. NATFW NSLP messages are initiated by the NSIS initiator (NI), handled by NSIS forwarders (NF) and finally processed by the NSIS responder (NR). It is required that at least NI and NR implement this NSLP, intermediate NF only implement this NSLP when they provide middlebox functions. Forwarders that do not have any NATFW NSLP functions just forward these messages; those forwarders implement NTLP and one or more other NSLPs.

A Data Sender (DS) that is intending to send data to a Data Receiver (DR) must start its NATFW NSLP signaling. So the NI at the data sender (DS) starts NSLP signaling towards the address of data receiver DR (see Figure 13).

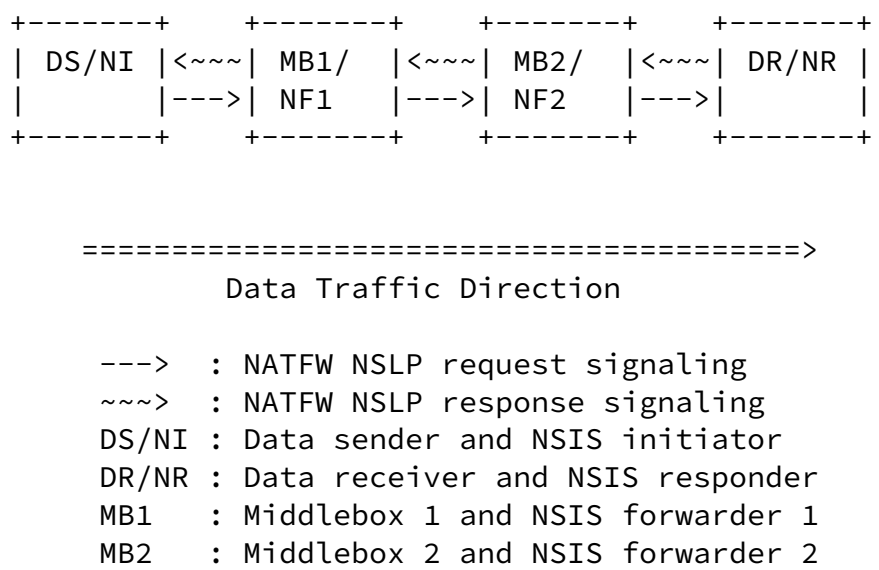


Figure 13: General NSIS signaling

The NSLP request messages are processed each time a NF with NATFW

NSLP support is passed. Those nodes process the message, check local policies for authorization and authentication, possibly create policy rules, and forward the signaling message to the next NSIS node. The request message is forwarded until it reaches the NSIS responder.

NSIS responders will check received messages and process those if applicable. NSIS responders generate response messages and sent them back to the NI via the same chain of NFs. The response message is processed at each NI forwarder implementing NATFW NSLP. The Data Sender can start sending its data flow to the Data Receiver, when the signaling was successful, meaning that NI has received a successful response.

In general, NATFW NSLP signaling follows the data path from DS to DR. This enables communication between both hosts for scenarios with only Firewalls on the data path or NATs on sender side. For scenarios with NATs on the receiver side certain problems arise, see also Section 2.

When Data receiver (DR) and Data Sender (DS) are located in different address realms and DR is behind a NAT, DS cannot signal to DR directly. DR is not reachable from DS and thus no NATFW signaling can be sent to DR's address. Therefore, DR must first determine an address at a NAT that is reachable for DS, for instance DR must determine its public IP address. Once DR has determined a public address it forwards this to DS via a separate mechanism, which may be application level signaling like SIP. This application level signaling may involve third parties that assist in exchanging this information. This separate mechanism is out of scope of NATFW NSLP.

NATFW NSLP signaling supports this public address fixing with this mechanism:

- First, DR determines a public address by signaling on the reverse path (DR towards DS) and thus making itself available to other hosts. This process of determining a public addresses is called reservation. This way DR reserves publicly reachable addresses and ports, but this address/port cannot be used by data traffic at this point of time.
- Second, DS is signaling directly to DR as DS would do if there is no NAT in between, and so creating policy rules at middleboxes. Note, that the reservation mode will make reservations only, which will be "activated" by the signaling from DS towards DR. The first mode is detailed in the Section 3.2.2

The protocol works on a soft-state basis, meaning that that whatever state is installed or reserved on a middlebox, it will expire, and thus be de-installed/ forgotten after a certain period of time. To prevent this, the involved boxes will have to specifically request a session extension. An explicit NATFW NSLP state deletion message is

also provided by the protocol.

Middleboxes should report back in case of error, so that appropriate measures and debugging can be performed.

The next sections define the NATFW NSLP message types and formats, protocol operations, and policy rule operations.

3.2 Protocol Operations

This section defines the protocol operations, how to create sessions, maintain them, and how to reserve addresses.

3.2.1 Creating Sessions

Allowing two hosts to exchange data even in the presence of middleboxes is realized in the NATFW NSLP by the 'create session' request message. The data sender generates a 'create session' message as defined in Section 3.4.2 and handles it to the NTLP. The NTLP forwards the whole message on the basis of the flow routing information towards DR. Each NSIS forwarders along the path that is implementing NATFW NSLP process the NSLP message, this is done NSLP hop-by-hop. Finally, the message is approaching DR, DR can accept the request or reject it. DR generates a response to the request, this response is transported hop by hop towards (XXX terminology) DS. NATFW NSLP forwarders may reject requests at any time. Figure 14 sketches the message flow between NI (DS), a NF (NAT), and NR (DR).

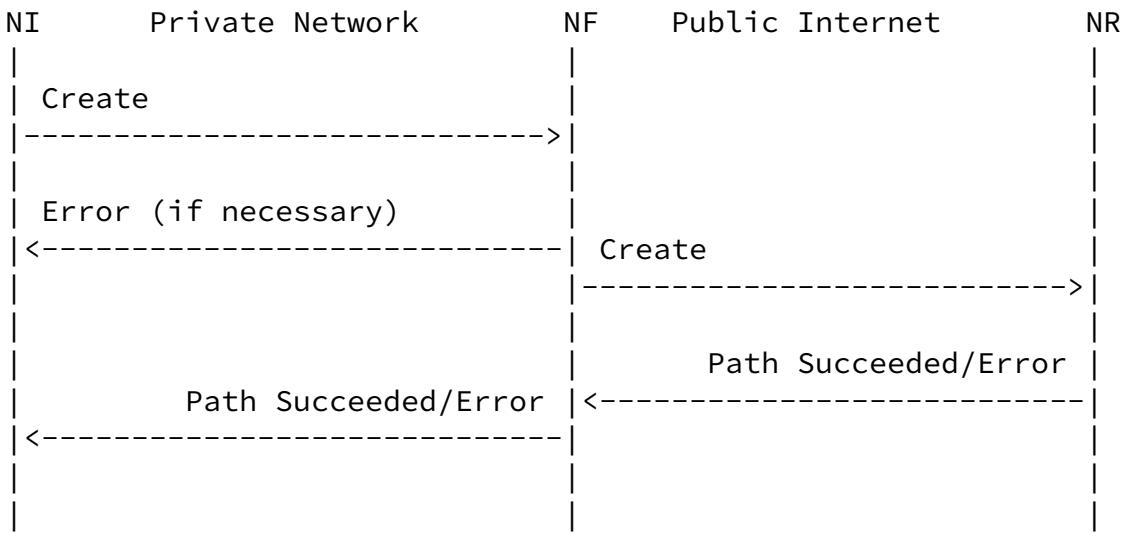


Figure 14: Creation message flow

Processing of 'create session' messages is differently per NSIS node:

- NSLP initiator: NI only generate 'create session' messages and handle them over to the NTLP. After receiving a 'path succeeded' the data path is configured and the NI can start sending its data to NR. After receiving an 'error' message the NI MAY try to generate the 'create session' message again or give up, depending on the error condition.
- NSLP forwarder: NSLP forwarders receiving 'create session' messages MUST first check authentication and authorization before any further processing is executed. The NF SHOULD check with its local policies if he can accept the desired policy rule given by NTLP's flow routing information. Further processing depends on the middlebox type:
  - NAT: When the 'create session' message is received at the public side a network external node is trying to open a NAT binding. First, it looks for a reservation made in advance by means of 'reserve external address' that matches the destination address/port of the flow routing information provided by the NTLP. If there is no reservation made in advance the NSLP SHOULD return an error message of type 'no reservation found' and discard the request. If there is a reservation, NSLP stores the data sender's address as part of the policy rule to be loaded and forwards the message with the address set to the internal address of the next NSIS node. When the 'create session' message is received at the private side the NAT binding is reserved, but not activated. The NSLP message is forwarded to next hop with source address set to the NAT's external address.
  - Firewall: When the 'create session' message is received the NSLP just remembers the requested policy rule, but does not install any policy rule. Afterwards, the message is forwarded to the next NSLP hop.
  - Combined NAT and Firewall: Processing at combined Firewall and NAT middleboxes is the same as in the NAT case. No policy rules are installed. Implementations MUST take care about the order of Firewall and NAT functions within the device. Order of functions is to be interpreted as how packets experience the treatment of those functions.
- NSLP receiver: NRs receiving 'create session' messages MUST reply with a 'path succeeded' message if they accept the request

message. Otherwise they SHOULD generate an error message. Both messages are sent back NSLP hop-by-hop towards NI.

Policy rules at middleboxes MUST be only installed upon receiving a successful response of type 'path succeeded'. This is a countermeasure to several problems, for instance, loaded policy rules at intermediate NF without reaching the actual NR.

3.2.2 Reserving External Addresses

NSIS signaling is intended to travel end-to-end, even in the presence of NATs and Firewalls on-path. This works well in cases where the data sender is itself behind a NAT and (covered by Section 3.2.1). For scenarios where the data receiver is located behind a NAT and it needs to receive data flows from outside its own network (see Figure 5) it is more troublesome. NSIS signaling, as well as subsequent data flows, are directed to a particular destination IP address that must be known in advance and reachable.

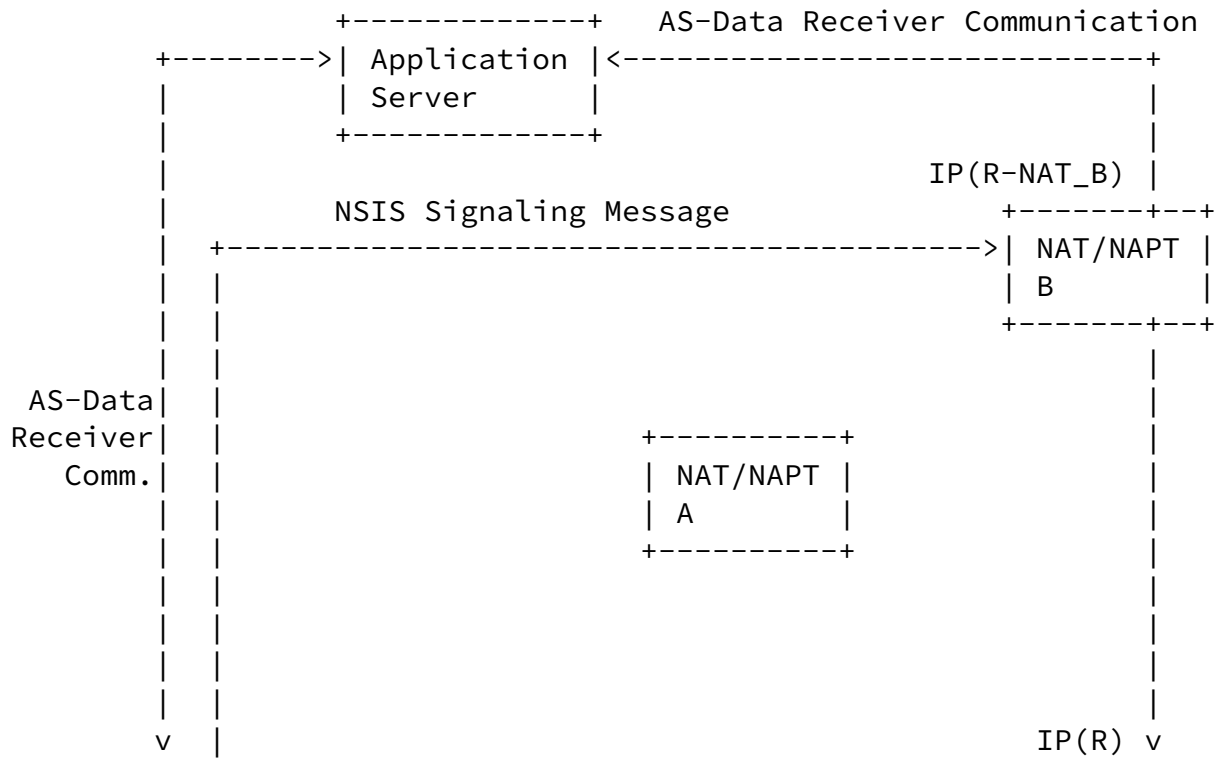


Figure 15: The Data Receiver behind NAT problem

Figure 15 describes a typical message communication in a peer-to-peer networking environment whereby the two end points learn of each others existence with the help of a third party (referred as Application Server). The communication with the application server and the two end points (data sender and data receivers) serves a number of functions. As one of the most important functions it

enables the two end hosts to learn the IP address of each other. The approach described in this memo supports this peer-to-peer approach, but is not limited to it.

Some sort of communication between the data sender/data receiver and a third party is typically necessary (independently of NSIS). NSIS signaling messages cannot be used to communicate application level relevant end point identifiers (in the generic case at least) as a replacement for the communication with the application server.

If the data receiver is behind a NAT then an NSIS signaling message will be addressed to the IP address allocated at the NAT (if there was one allocated). If no corresponding NSIS NAT Forwarding State at NAT/NAPT B exists (binding IP(R-NAT B) <-> IP(R)) then the signaling message will terminate at the NAT device (most likely without proper response message). The signaling message transmitted by the data sender cannot install the NAT binding or NSIS NAT Forwarding State "on-the-fly" since this would assume that the data sender knows the topology at the data receiver side (i.e. the number and the arrangement of the NAT and the private IP address(es) of the data receiver). The primary goal of path-coupled middlebox communication was not to force end hosts to have this type of topology knowledge.

Public Internet Private Address

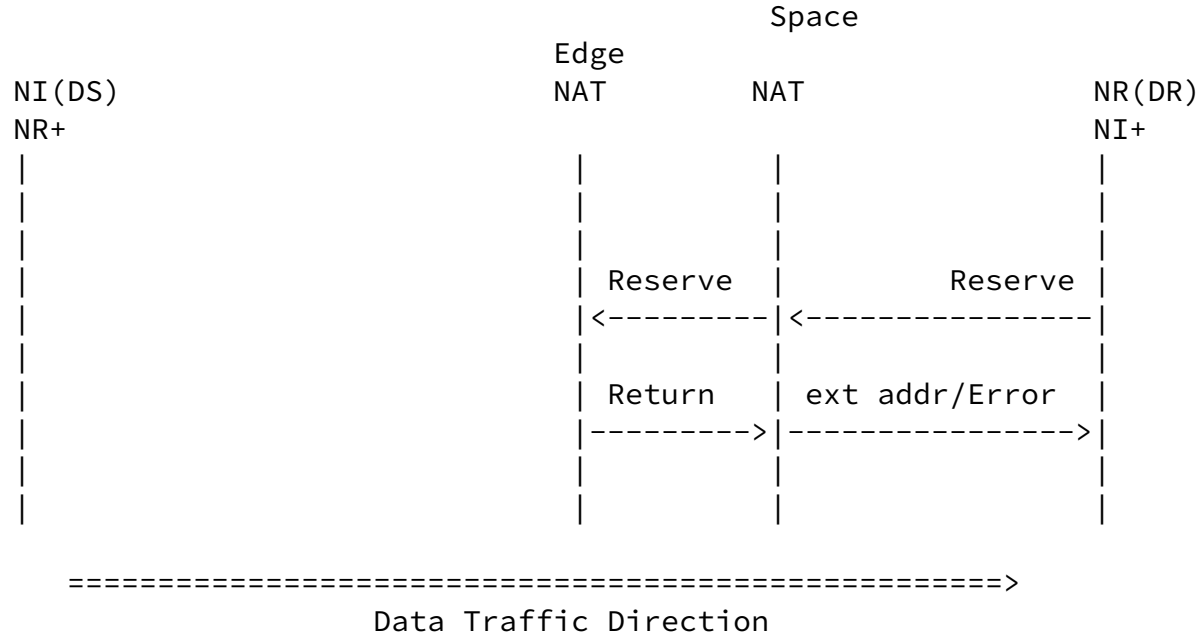


Figure 16: Reservation message flow

Figure 16 shows the message flow for reserving an external address/ port at a NAT. In this case the roles of the different NSIS entities are:

- The actual data receiver (DR) is the NSIS initiator (NI+) for the 'reserved external address' message, but the NSIS responder (NR) for 'create session' messages following later.
- The actual data sender (DS) will be the NSIS initiator (NI) for later 'create session' messages and may be the NSIS target of the signaling (NR+).
- The actual target of the 'reserved external address' message may be an arbitrary address NR+.

The data receiver DR starts to signal an 'reserve external address' message into the "wrong direction". By "wrong" we refer to the usual behavior of path-coupled signaling where the data sender starts signaling in order to tackle with routing asymmetry. The data receiver would typically return signaling messages to the data sender in the reverse direction by utilizing state created at nodes along the path (i.e. to reverse route signaling messages). In case of

establishing NAT bindings (and NSIS NAT Forwarding State) the direction does not matter since the data path is modified through route pinning due to the external NAT address. Subsequent NSIS messages (and also data traffic) will travel through the same NAT boxes. The signaling target address selection for this message is discussed in Section 3.2.10.

The signaling message creates NSIS NAT Forwarding State at intermediate NSIS NAT node(s). Furthermore it has to be ensured that the edge NAT device is discovered as part of this process. The end host cannot be assumed to know this device - instead the NAT box itself is assumed to know that it has such a capability. Forwarding of the 'reserve external address' message beyond this entity is not necessary, and should be prohibited as it provides information on internal hosts capabilities.

The edge NAT device is responding with a 'return external address' message containing the public reachable IP address/port number.

Processing of 'reserve external address' messages is differently per NSIS node:

- NSLP initiator: NI+ only generate 'reserve external address' messages and should never receive them.
- NSLP forwarder: NSLP forwarders receiving 'reserve external address' messages MUST first check authentication and authorization before any further processing is executed. The NF SHOULD check with its local policies if he can accept the desired policy rule given by NTLP's flow routing information. Further processing depends on the middlebox type:

  * NAT: NATs check whether the message is received at the public address or at the private address. If received at the public address a NF MAY generate an error message of type 'requested external address from outside'. If received at the private address, an IP address/port is reserved. In the case it is an edge-NAT, the NSLP message is not forwarded anymore and a response of type 'return external address' is generated. If it is not an edge-NAT, the NSLP message is forwarded further.
  * Firewall: Firewalls MUST not change their configuration upon a 'reserve external address' message. They simply MUST forward the message and MUST keep NTLP state. Firewalls that are configured as edge-Firewalls (XXX, do definition!) MAY return

an error of type 'no NAT here'.
    * Combined NAT and Firewall: Processing at combined Firewall and NAT middleboxes is the same as in the NAT case.
- NSLP receiver: This type of message should never be received by any NR and it SHOULD be discarded silently.

Processing of 'return external address' messages is differently per NSIS node:
- NSLP initiator: Upon receiving a 'return external address' message the NI+ can use the obtained IP address and port number for further application signaling.
- NSLP forwarder: NFs simply forward this message as long as they keep state for the requested reservation.
- NSIS responder: This type of message should never be received by any NR and it SHOULD be discarded silently.

3.2.3 Reserving External Addresses and Create Session

Some migration scenarios need specialized support to cope with the situation where the receiving side is running NSIS only. End-to-end signaling is going to fail without NSIS support at both sides. For this the 'create-reverse' signaling mode is supported. In this case, a DR can signal towards the DS like in the 'reserve external address' message scenario. The message is forwarded until it reaches the edge-NAT and retrieves a public IP address and port number. Unlike in the 'reserve external address' no 'return external address' response message is created, the forwarding of the request message stops and a 'create session' message is generated by the edge-NAT. This request message is sent towards DR with DS as source address and follows the regular processing orders as 'create session' messages do. The exact definition of this mode is to be done.

3.2.4 Prolonging Sessions

NATFW NSLP sessions are maintained on a soft-state base. After a certain timeout sessions and corresponding policy rules are removed

automatically by the middlebox, if they are not refreshed by a prolong session message. NI is sending prolong message towards NR and each NSIS forwarder maintaining state for the given session ID extends the lifetime of the session. Extending lifetime of a session is calculated as current local time plus lifetime. Section 3.2.7 is

defining the process of calculating lifetimes in detail.

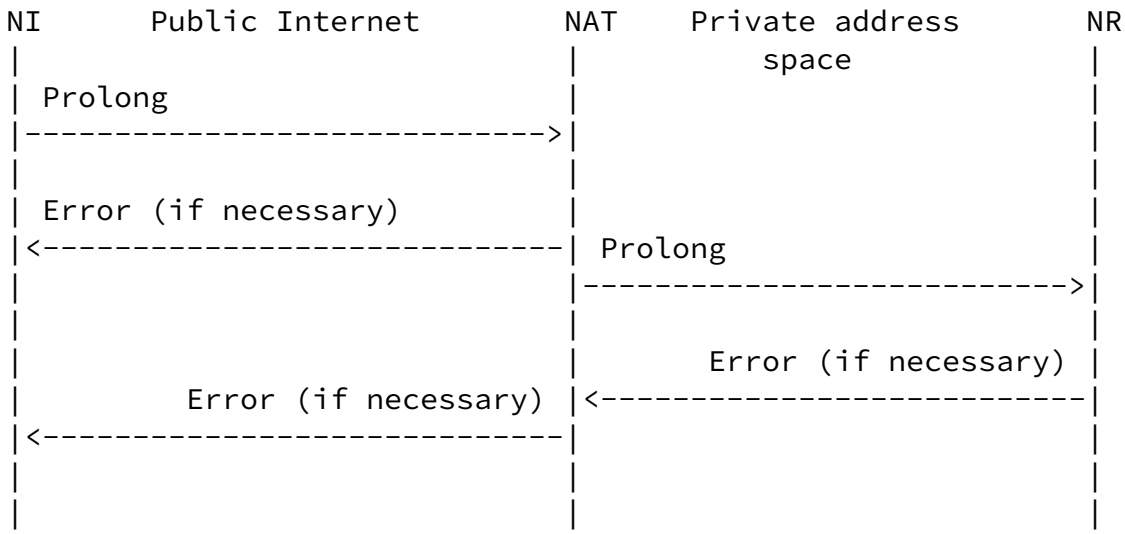


Figure 17: Prolongation message flow

Processing of 'prolong session' messages is differently per NSIS node:

- NSLP initiator: NI can generate 'prolong session' messages before the session times out.
- NSLP forwarder: NSLP forwarders receiving 'prolong session' messages MUST first check authentication and authorization before any further processing is executed. The NF SHOULD check with its local policies if he can accept the desired lifetime extension for the session referred by the session ID. Processing of this message is independent of the middlebox type.
- NSLP responder: NIs accepting this prolong message generate a 'path succeeded' message.

3.2.5 Deleting Sessions

NATFW NSLP sessions may be deleted at any time. NSLP initiators can trigger this deletion via the 'delete session' message, as shown in Figure 17.

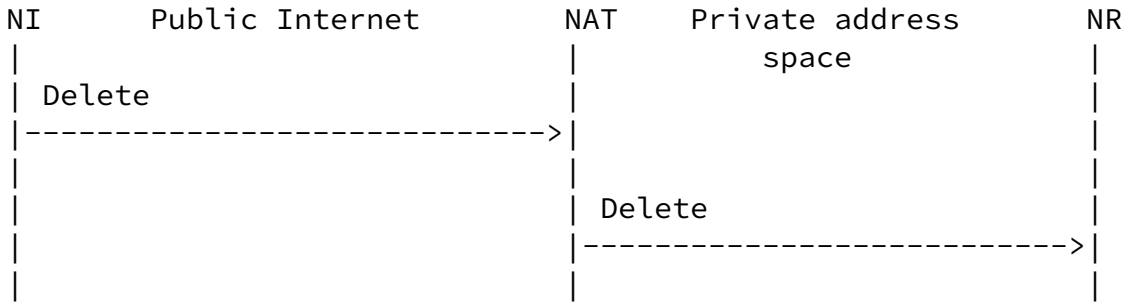


Figure 18: Delete message flow

NSLP nodes receiving this message MUST delete the session immediately. Corresponding policy rules to this particular session MUST be deleted immediately, too. This message is forwarded until it reaches the final NR. The 'delete' message does not generate any response, neither positive nor negative, since there is no NSIS state left at the nodes along the path.

3.2.6 Authorization

Authorization and security issues are currently discussed in a different document and will be included after reaching consensus ([20]).

3.2.7 Calculation of Lifetimes

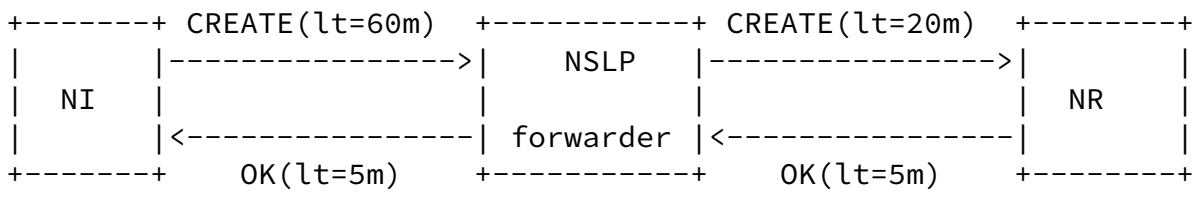
NATFW NSLP sessions, and the corresponding policy rules possibly installed, are maintained via soft-state. Each session is assigned a lifetime and they are kept alive as long as the lifetime is valid. After the expiration of the lifetime sessions and policy rules MUST be removed automatically and resources bound to them should be freed as well. Session lifetime is kept at every NATFW NSLP node. The NSLP forwarders and NSLP responder are not responsible for triggering lifetime prolongination messages (see Section 3.2.4), this is the task of the NSIS initiator.

NSIS initiator MUST choose a lifetime value before they can sent any message (except 'delete session' messages) to other NSLP nodes. This lifetime value should consider application's needs, i.e., duration in terms of minutes or hours, and networking needs, i.e., values in the range less than 30 seconds may not be useful. This requested lifetime value is placed in the 'lifetime object' of the NSLP message and messages are forwarded to the next NATFW NSLP node.

NATFW NSLP forwarders processing the request message along the path MAY lower the request lifetime given to fit their needs and/or local

policy. NATFW forwarders MUST NOT increase the lifetime value; they MAY reject the requested lifetime immediately and MUST generate an error response message of type 'lifetime too big' upon rejection. The NSLP request message is forwarded until it reaches the NSLP responder. NSLP responder MAY reject the requested lifetime value and MUST generate an error response message of type 'lifetime too big' upon rejection. NSLP responder MAY lower the requested lifetime as well to a granted lifetime. NSLP responders generate their appropriate response message for the received request message, sets the lifetime value to the above granted lifetime and sends the message back hop-by-hop towards NSLP initiator.

Each NSLP forwarder processes the response message, reads and stores the granted lifetime value. The forwarders SHOULD accept the granted lifetime, as long as the value is equal or lower than the requested lifetime. They MAY reject the lifetime and generate a 'lifetime not acceptable' error response message. Figure 19 shows the procedure with an example, where an initiator requests 60 minutes lifetime in 'create session' message and the lifetime is shortened along the path by the forwarder to 20 minutes and by the responder to 5 minutes.



lt = lifetime
CREATE = 'create session' message
OK = 'path succeeded' message

Figure 19: Lifetime Calculation Example

3.2.8 Middlebox Resource

This section needs to be done and should describe how to map flow

routing information to middlebox policy rules. Further, this section should clarify wildcarding. XXX

3.2.9 De-Multiplexing at NATs

Section 3.2.2 describes how NSIS nodes behind NATs can obtain a

public reachable IP address and port number at a NAT. The information IP address/port number can be transmitted via a signaling protocol and/or third party to the communication partner that would like to send data towards. However, NSIS signaling flows are sent towards the address of the NAT at which this particular IP address and port number is allocated. The NATFW NSLP forwarder at this NAT needs to know how the incoming NSLP requests are related to reserved addresses, meaning how to de-multiplex incoming requests.

Two options for de-multiplexing incoming NSLP requests are:
1. Based on flow routing information, like protocol number and TCP port numbers.
2. Based on NSIS session IDs.

Approach 2) would require that both NSIS ends, initiator and responder, use the same session ID in NSIS signaling. Since session IDs are usually generated randomly, application level signaling would have to be adapted to carry NSIS session IDs used during reservation to the other end (the NSIS initiator sending the 'create session' message). This approach SHOULD NOT be used.

Approach 1) uses information stored at NATs (like mapping of public IP address to private, transport protocol, port numbers) and information given by NTLP's flow routing information to de-multiplex NSIS messages. This approach is RECOMMENDED.

3.2.10 Selecting Destination IP addresses for REA

Request messages of type 'reserve external address' do need, as any other message type as well, a final destination IP address to reach. But as many applications do not provide a destination IP address at the first place, there is a need to choose a destination address for the 'reserve external address' messages. This destination can be the final target, but for the mentioned type of application, the destination address can be arbitrary. Taking the "correct"

destination IP address might be difficult and there is no right answer. [19] shows choices for SIP and this section provides some hints about choosing a good destination IP address in general.

1. Public IP address of the data sender:
   * Assumption:
     + The data receiver already learned the IP address of the data sender (e.g. via a third party).
   * Problems:
     + The data sender might also be behind a NAT. In this case the public IP address of the data receiver is the IP address allocated at this NAT.

     + Due to routing asymmetry it might be possible that the routes taken by a) the data sender and the application server b) the data sender and NAT B might be different. As a consequence it might be necessary to advertise a new (and different) external IP address with SIP after using NSIS to establish a NAT binding.
2. Public IP address of the data receiver (allocated at NAT B):
   * Assumption:
     + The data receiver already learned his externally visible IP address (e.g. based on the third party communication).
   * Problems:
     + Communication with a third party is required.
3. IP address at the Application Server:
   * Assumption:
     + An application server (or a different third party) is available.
   * Problems:
     + If the NSIS signaling message is not terminated at the NAT of the local network then an NSIS unaware application server might discard the message.
     + Routing might not be optimal since the route between a) the data receiver and the application server b) the data receiver and the data sender might be different.

3.3 NATFW NSLP Messages Components

A NATFW NSLP message consists of a NSLP header and one or more objects following the header. The NSLP header is common for all

NSLPs and objects are Type-Length-Value (TLV) encoded using big endian (network ordered) binary data representations. Header and objects are bound to 32 bits and objects that do not fall into 32 bits boundaries must be padded to 32 bits.

The whole NSLP message is carried in a NTLP message.

Note that the notation 0x is used to indicate hexadecimal numbers.

3.3.1 NSLP Header

The NSLP header is common to all NSLPs and is the first part of all NSLP messages. It contains two fields, the NSLP message type and a reserved field. The total length is 32 bits. The layout of the NSLP header is defined by Figure 20.

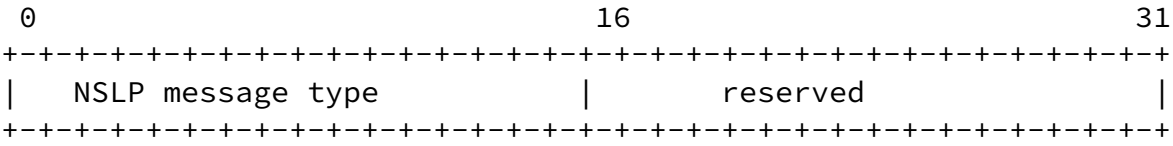


Figure 20: Common NSLP header

The reserved field MUST be set to zero in the NATFW NSLP header before sending and MUST be ignored during processing the header. Note that other NSLPs use this field as flag field.

3.3.2 NSLP message types

The message types identify requests and responses. Defined messages types for requests are:

- 0x0101 : create
- 0x0102 : reserve
- 0x0103 : reserve-create
- 0x0104 : prolong

o 0x0105 : delete

Defined message types for responses are:

o 0x0201 : path_succeed
o 0x0202 : path_deleted
o 0x0203 : ret_ext_addr
o 0x0204 : error

3.3.3 NSLP Objects

NATFW NSLP objects use a common header format defined by Figure 21. Objects are Type-Length-Value (TLV) encoded using big endian (network ordered) binary data representations. The object header contains two fields, the NSLP object type and the object length. Its total length is 32 bits.

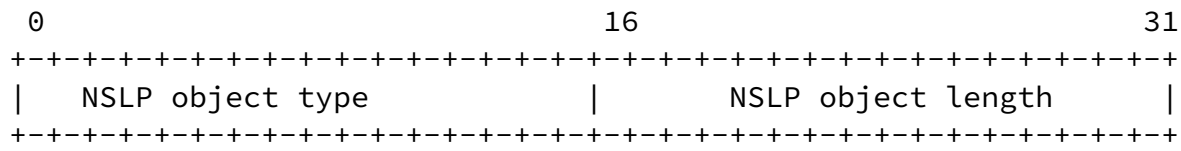


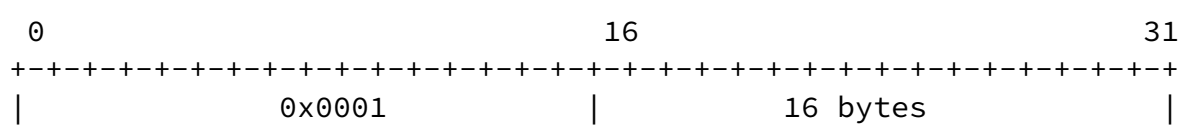
Figure 21: Common NSLP object header

The length is the total length of the object without the object

header. The unit is bytes. The particular values of type and length for each NSLP object are listed in the subsequent chapters that define the NSLP objects.

3.3.3.1 Session ID Object

The session ID object carries an identifier for the session of the signaled flow. The only field is the session ID of 16 bytes length.



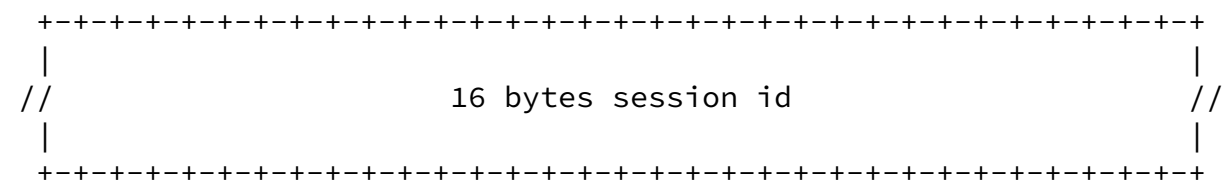


Figure 22: Session ID object

The session ID is generated in random way by the NSIS initiator.

3.3.3.2 Session Lifetime Object

The session lifetime object carries the requested or granted lifetime of a NATFW NSLP session measured in seconds. The object consists only of the 4 bytes lifetime field.

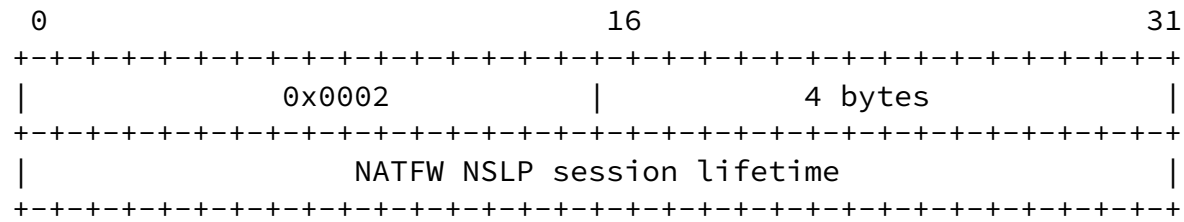


Figure 23: Lifetime object

3.3.3.3 External Address Object

The external address objects can be included in ret_ext_addr responses (Section 3.4.9) only. It contains the external IP address and port number allocated at the edge-NAT. Note that this address/port may be either reserved or reserve-create. Two fields are defined, the external IP address, and the external port number. For

IPv4 the object with value 0x0010 is defined. It has a length of 8 bytes and is shown in Figure 24.

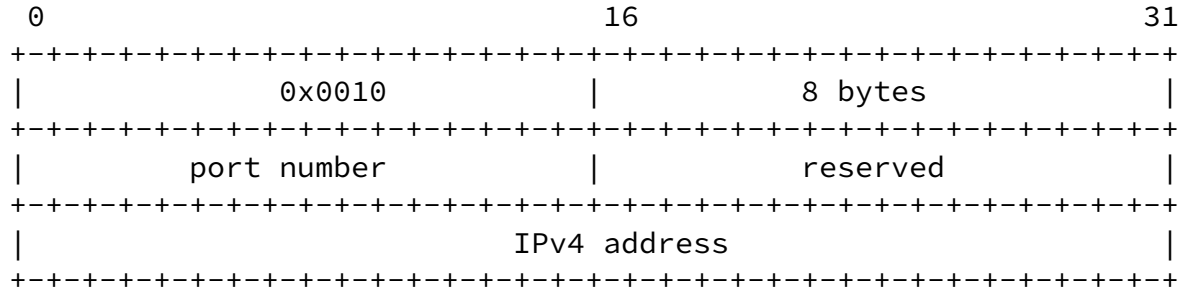


Figure 24: External Address Object for IPv4 addresses

For IPv6 the object with value 0x0011 is defined. It has a length of 20 bytes and is shown in Figure 25.

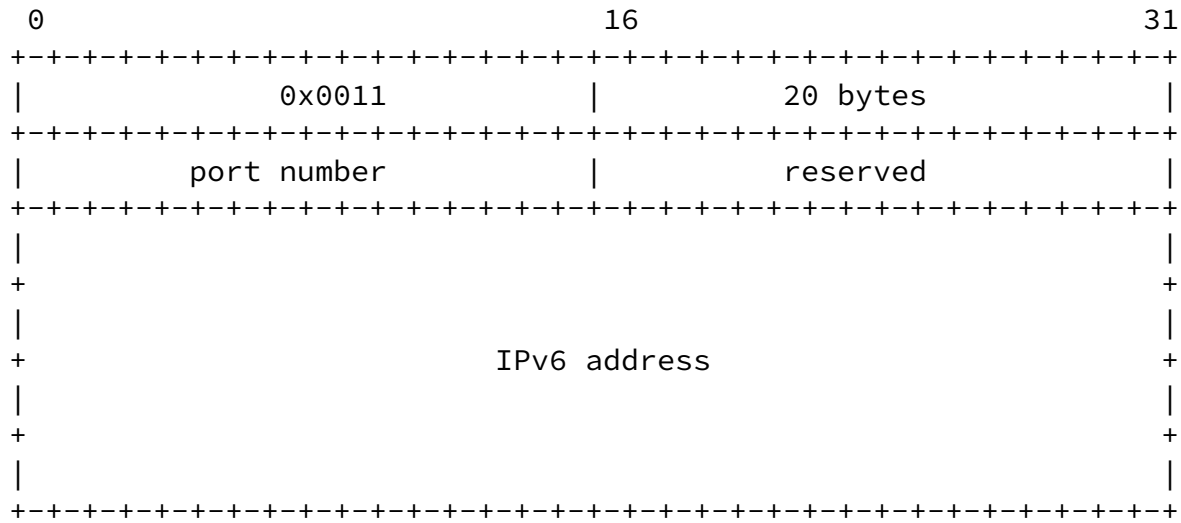


Figure 25: External Address Object for IPv6 addresses

3.3.3.4 Extended Flow Information Object

In general, flow information is kept at the NTLP level during signaling. Nevertheless, some additional information may be required for NSLP operations. The 'extended flow information' object carries this additional information about number of subsequent port numbers that should be allocated at middleboxes.

These fields are defined for the policy rule object:

- Number of ports: This field gives the number of ports that should be allocated beginnig at the port given in NTLP's flow routing information.

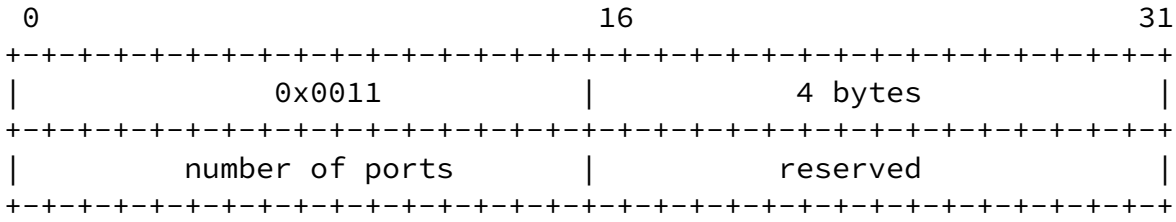


Figure 26: Extended Flow Information

3.3.3.5 Error Object

The error object carries the reason for an error. It has only one field, the error code, and is 2 bytes long.

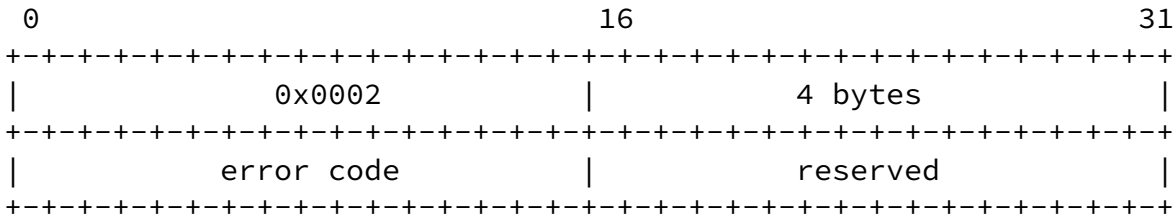


Figure 27: Error

TBD: Define error clases and define the error coded. Possible classes are:

- Policy rule errors
- Authentication and Authorization errors
- NAT

Currently in this memo defined errors:

- lifetime too big
- lifetime not acceptable
- no NAT here
- no reservation found
- requested external address from outside

3.4 Message Formats

This section defines the content of each NATFW NSLP message type. The message types are defined in Section 3.3.2. First, the request messages are defined with their respective objects to be included in the message. Second, the response messages are defined with their respective objects to be included.

Basically, each message is constructed of NSLP header and one or more NSLP objects. The order of objects is not defined, meaning that objects may occur in any sequence.

Each section elaborates the required settings and parameters to be set by the NSLP at the NTLP, for instance, how the flow routing information is set.

3.4.1 Policy Rules

Policy rules are the building block of middlebox devices considered in the NATFW NSLP. For Firewalls the policy rule consists usually of a 5-tuple, source/destination address, transport protocol, and source/destination port number, plus an action like allow or deny. Other actions are available depending on the implementation of the Firewall, but NATFW NSLP uses only allow action, since a default to deny policy at the middlebox is assumed. For NATs the policy rule consists of action 'map this another address realm' and further mapping information, that might be in the most simply case internal IP address and external IP address.

Policy rules are usually carried in one piece in signaling applications. In NSIS the policy rule is divided into the filter specification, an implicit allow action, and additional information. The filter specification is carried within NTLP's flow routing

information and additional information is carried in NSLP's objects. Additional information is for instance the lifetime of a policy rule or session.

3.4.2 Create Session (CRS)

The create session request message is used to create NSLP sessions and at middleboxes to create policy rules.

The create session messages carries these objects:
- Session ID object
- Lifetime object

The flow routing information in the NTLP MUST be set to DS as source address and DR as destination address. All other parameters MUST be set according the required policy rule.

3.4.3 Reserve External Address (REA)

The reserve external address (REA) request message is used to lookup a NAT and to allocated an external IP address and possibly port number, so that the initiator of the REA request has a public reachable IP address/port number.

The REA request message carries these objects:
- Session ID object
- Lifetime object

The REA message needs special NTLP treatment. First of all, REA messages travel the wrong way, from the DR towards DS. Second, the DS' address used during the signaling may be not the actual DS (see Section 3.2.10). Therefore, the NTLP flow routing information is set to DR as initiator and DS as responders, a special field is given in the NTLP: The signaling destination.

3.4.4 Reserve-Create (REC)

XXX This is a proposal for a new message to support the reservation and simultaneous/implicit create message generation.

The reserve-create message carries these objects:
o Session ID object
o Lifetime object

NTLP issues: TBD.

3.4.5 Prolong Session (PLS)

The prolong request message is used to prolong (extend) the lifetime of a NATFW NSLP and policy rules at middleboxes.

The prolong session message carries these objects:

o Session ID object
o Lifetime object

The flow routing information in the NTLP MUST be set to DS as source address and DR as destination address. All other parameters MUST be set according the required policy rule.

3.4.6 Delete Session (DLS)

The delete request message is used to delete NATFW NSLP sessions.

The delete session message carries these objects:
o Session ID object

The flow routing information in the NTLP MUST be set to DS as source address and DR as destination address. All other parameters MUST be set according the required policy rule.

3.4.7 Path Succeeded (PS)

The path succeeded response message is used to acknowledge a successful create and prolong.

The path succeeded message carries these objects:
o Session ID object
o lifetime object

This message is routed on the reverse path.

3.4.8 Path Deleted (PD)

The path deleted response message is used to acknowledge a successful delete request message.

The path deleted message carries this object:
- Session ID object

This message is routed on the reverse path.

3.4.9 Return External Address (RA)

The return external address response message is sent back as a positive result of reserve external address request. It contains the reserved external IP address and port number.

The path succeeded message carries these objects:
- Session ID object

- Lifetime object
- External address object (either IPv4 or IPv6 type)

This message is routed on the reverse path.

3.4.10 Error Response (ER)

The error response message is sent back by any NSIS node involved in the session that occurs an error condition.

The error message carries these objects:
- Session ID object
- Error object

This message is routed on the reverse path.

4. NSIS NAT and Firewall transitions issues

NSIS NAT and Firewall transition issues are premature and will be addressed in a separate draft (see [17]). An update of this section will be based on consensus.

5. Security Considerations

Security is of major concern particularly in case of Firewall traversal. Generic threats for NSIS signaling have been discussed in [6] and are applicable here as well. It is necessary to provide proper signaling message protection and proper authorization. Note that the NAT is likely to be co-located with a Firewall and might therefore require packet filters to be changed in order to allow the

signaling message to process and to traverse. This section aims to raise some items for further discussion and illustrates the problems the authors faced when creating a security solution for the NAT/ Firewall NSLP.

Installing packet filters provides some security, but has some weaknesses, which heavily depend on the type of packet filter installed. A packet filter cannot prevent an adversary to inject traffic (due to the IP spoofing capabilities). This type of attack might not be particular helpful if the packet filter is a standard 5 tuple which is very restrictive. If packet filter installation, however, allows specifying a rule, which restricts only the source IP address, then IP spoofing allows transmitting traffic to an arbitrary address. NSIS aims to provide path-coupled signaling and therefore an adversary is somewhat restricted in the location from which attacks can be performed. Some trust is therefore assumed from nodes and networks along the path.

Without doubts there is a dependency on the security provided by the NTLP. Section Appendix A and Section 2.2 motivates some trust relationship and authorization scenarios. These scenarios deserve a discussion since some of them (particularly one with a missing network-to-network trust relationship) is different to what is know from QoS signaling. To address some of these trust relationships and authorization issues requires security mechanisms between non-neighboring nodes at the NSLP layer. For the group of authors it seems that peer-to-peer and end-to-middle security needs to be provided. An NSLP security mechanism between neighboring NSLP peers might be necessary if security mechanisms at the NTLP do not provide adequate protection mechanisms. This issue is, however, still in discussion.

As a design goal it seems to be favorable to reuse existing mechanisms to the best extend possible. In most cases it is necessary to carry the objects for end-to-middle as NSLP payloads since the presence of NATs might prevent direct communication. Three security mechanisms have to be considered in more detail in a future version of this document: CMS [21] and Identity Representation for RSVP [15]. The authors believe that CMS more suitable (since it provides much more functionality). The details deserve further

discussion and implementation experience.

With regard to signal between two end hosts even though the receiver is behind a NAT this proposal suggests creating state by the data receiver first. This allows NSIS signaling messages to traverse a NAT at the receiver side (due to the established state at this NAT box) and simplifies security handling. To achieve this behavior it is required to install NSIS NTLP and NSLP state. Furthermore, it is envisioned to associate the two signaling parts (one part from the data sender to the NAT and the other part from the NAT to the data receiver) with the help of the Session Identifier. As such, the discussion in [15] is relevant for this document.

Another interesting property of this protocol proposal is to prevent Denial of Service attacks against NAT boxes whereby an adversary allocates NAT bindings with the help of data packets. Since these data packets do not provide any type of authentication and are not authorized any adversary is able to mount such an attack. This attack has been mentioned at several places in the literature already and is particularly harmful if no NAPT functionality is used (i.e. if a new NAT binding consumes one IP address of a pool of IP addresses). Using the protocol described in this document additional security can be achieved and more fairness can be provided.

6. Open Issues

At least the following issues require further discussion:

- Option processing rules in presence of unknown options.
- Terminology w.r.t. the term wrong way.
- NTLP: New object and semantics for REA.
- NTLP and NATFW NSLP interaction
- List of NTLP transport modes per NSLP message
- Routing Change detection
- Query message, definition of semantics needed
- Is there a need for a QoS NSLP RSN like object/mechanism in NATFW NSLP?
- Add IANA considerations section.
- re-work security considerations.
- Query message: Syntax and semantics.
- Add text about asynchronous messages.
- Anycast address for REA.
- Check common formats with QoS NSLP
- Change length field of objects to long words as unit?
- Variable length for session id?
- Meaning of 0 as session ID.
- Extended flow object: Needs refinement

7. Contributors

A number of individuals have contributed to this draft. Since it was not possible to list them all in the authors section, it was decided to split it and move Marcus Brunner and Henning Schulzrinne into the contributors section. Separating into two groups was done without treating any one of them better (or worse) than others.

8. References

8.1 Normative References

[1] Hancock et al, R., "Next Steps in Signaling: Framework", DRAFT draft-ietf-nsis-fw-05.txt, October 2003.

[2] Brunner et al., M., "Requirements for Signaling Protocols", DRAFT draft-ietf-nsis-req-09.txt, October 2003.

[3] Schulzrinne, H. and R. Hancock, "GIMPS: General Internet Messaging Protocol for Signaling", DRAFT draft-ietf-nsis-ntlp-00.txt, October 2003.

[4] Van den Bosch, S., Karagiannis, G. and A. McDonald, "NSLP for Quality-of-Service signaling", DRAFT draft-ietf-nsis-qos-nslp-03.txt, May 2004.

[5] IANA, "Special-Use IPv4 Addresses", RFC 3330, September 2002.

[6] Tschofenig, H. and D. Kroeselberg, "Security Threats for NSIS", DRAFT draft-ietf-nsis-threats-01.txt, January 2003.

[7] Srisuresh, P., Kuthan, J., Rosenberg, J., Molitor, A. and A. Rayhan, "Middlebox communication architecture and framework", RFC 3303, August 2002.

8.2 Informative References

[8] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations, RFC 2663", August 1999.

[9] Srisuresh, P. and M. Holdrege, "Network Address Translator (NAT)Terminology and Considerations, RFC 2663".

[10] Srisuresh, P. and E. Egevang, "Traditional IP Network Address Translator (Traditional NAT), RFC 3022", January 2001.

[11] Tsirtsis, G. and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT), RFC 2766", February 2000.

[12] Carpenter, B. and S. Brim, "Middleboxes: Taxonomy and Issues", RFC 3234, February 2002.

[13] Srisuresh, P., Tsirtsis, G., Akkiraju, P. and A. Heffernan, "DNS extensions to Network Address Translators (DNS_ALG)", RFC 2694, September 1999.

[14] Braden, B., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", September 1997.

[15] Yadav, S., Yavatkar, R., Pabbati, R., Ford, P., Moore, T., Herzog, S. and R. Hess, "Identity Representation for RSVP", RFC 3182, October 2001.

[16] Tschofenig, H., Schulzrinne, H., Hancock, R., McDonald, A. and X. Fu, "Security Implications of the Session Identifier", June 2003.

[17] Aoun, C., Brunner, M., Stiemerling, M., Martin, M. and H. Tschofenig, "NAT/Firewall NSLP Migration Considerations", DRAFT draft-aoun-nsis-nslp-natfw-migration-01.txt, Februar 2004.

[18] Aoun, C., Brunner, M., Stiemerling, M., Martin, M. and H. Tschofenig, "NATFirewall NSLP Intra-realm considerations", DRAFT draft-aoun-nsis-nslp-natfw-intrarealm-00.txt, Februar 2004.

[19] Martin, M., Brunner, M. and M. Stiemerling, "SIP NSIS Interactions for NAT/Firewall Traversal", DRAFT draft-martin-nsis-nslp-natfw-sip-00.txt, Februar 2004.

[20] Martin, M., Brunner, M., Stiemerling, M., Girao, J. and C. Aoun, "A NSIS NAT/Firewall NSLP Security Infrastructure", DRAFT draft-martin-nsis-nslp-natfw-security-01.txt, Februar 2004.

[21] Housley, R., "Cryptographic Message Syntax (CMS)", RFC 3369, August 2002.

[22] Manner, J., Suihko, T., Kojo, M., Liljeberg, M. and K. Raatikainen, "Localized RSVP", DRAFT draft-manner-lrsvp-00.txt, November 2002.

[23] Tschofenig, H., Buechli, M., Van den Bosch, S. and H. Schulzrinne, "NSIS Authentication, Authorization and Accounting Issues", March 2003.

[24] Amini, L. and H. Schulzrinne, "Observations from router-level internet traces", DIMACS Workshop on Internet and WWW Measurement, Mapping and Modelin Jersey) , Februar 2002.

[25] Adrangi, F. and H. Levkowetz, "Problem Statement: Mobile IPv4 Traversal of VPN Gateways", draft-ietf-mobileip-vpn-problem-statement-req-02.txt (work in progress), April 2003.

[26] Ohba, Y., Das, S., Patil, P., Soliman, H. and A. Yegin, "Problem Space and Usage Scenarios for PANA", draft-ietf-pana-usage-scenarios-06 (work in progress), April 2003.

[27] Shore, M., "The TIST (Topology-Insensitive Service Traversal) Protocol", DRAFT draft-shore-tist-prot-00.txt, May 2002.

[28] Tschofenig, H., Schulzrinne, H. and C. Aoun, "A Firewall/NAT Traversal Client for CASP", DRAFT draft-tschofenig-nsis-casp-midcom-01.txt, March 2003.

[29] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M. and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[30] Brunner, M., Stiemerling, M., Martin, M., Tschofenig, H. and H. Schulzrinne, "NSIS NAT/FW NSLP: Problem Statement and

Framework", DRAFT draft-brunner-nsis-midcom-ps-00.txt, June 2003.

[31] Ford, B., Srisuresh, P. and D. Kegel, "Peer-to-Peer(P2P) communication Network Address Translators(NAT)", DRAFT draft-ford-midcom-p2p-02.txt, March 2004.

[32] Rosenberg et al, J., "STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)", RFC 3489, March 2003.

[33] Rekhter et al, Y., "Address Allocation for Private Internets", RFC 1918, February 1996.

Authors' Addresses

Martin Stiemerling
Network Laboratories, NEC Europe Ltd.
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Phone: +49 (0) 6221 905 11 13
EMail: stiemerling@netlab.nec.de
URI:

Hannes Tschoefenig
Siemens AG
Otto-Hahn-Ring 6
Munich 81739
Germany

Phone:
EMail: Hannes.Tschofenig@siemens.com
URI:

Miquel Martin
Network Laboratories, NEC Europe Ltd.
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Phone: +49 (0) 6221 905 11 16
EMail: miquel.martin@netlab.nec.de
URI:

Cedric Aoun
Nortel Networks

France

EMail: cedric.aoun@nortelnetworks.com

Appendix A. Problems and Challenges

This section describes a number of problems that have to be addressed for NSIS NAT/Firewall. Issues presented here are subject to further

discussions. These issues might be also of relevance to other NSLP protocols.

A.1 Missing Network-to-Network Trust Relationship

Peer-to-peer trust relationship, as shown in Figure 10, is a very convenient assumption that allows simplified signaling message processing. However, it might not always be applicable, especially between two arbitrary access networks (over a core network where signaling messages are not interpreted). Possibly peer-to-peer trust relationship does not exist because of the large number of networks and the unwillingness of administrators to have other network operators to create holes in their Firewalls without proper authorization. Hence in the following scenario we assume a somewhat different message processing and show three possible approaches to tackle the problem. None of these three approaches is without drawbacks or constraining assumptions.

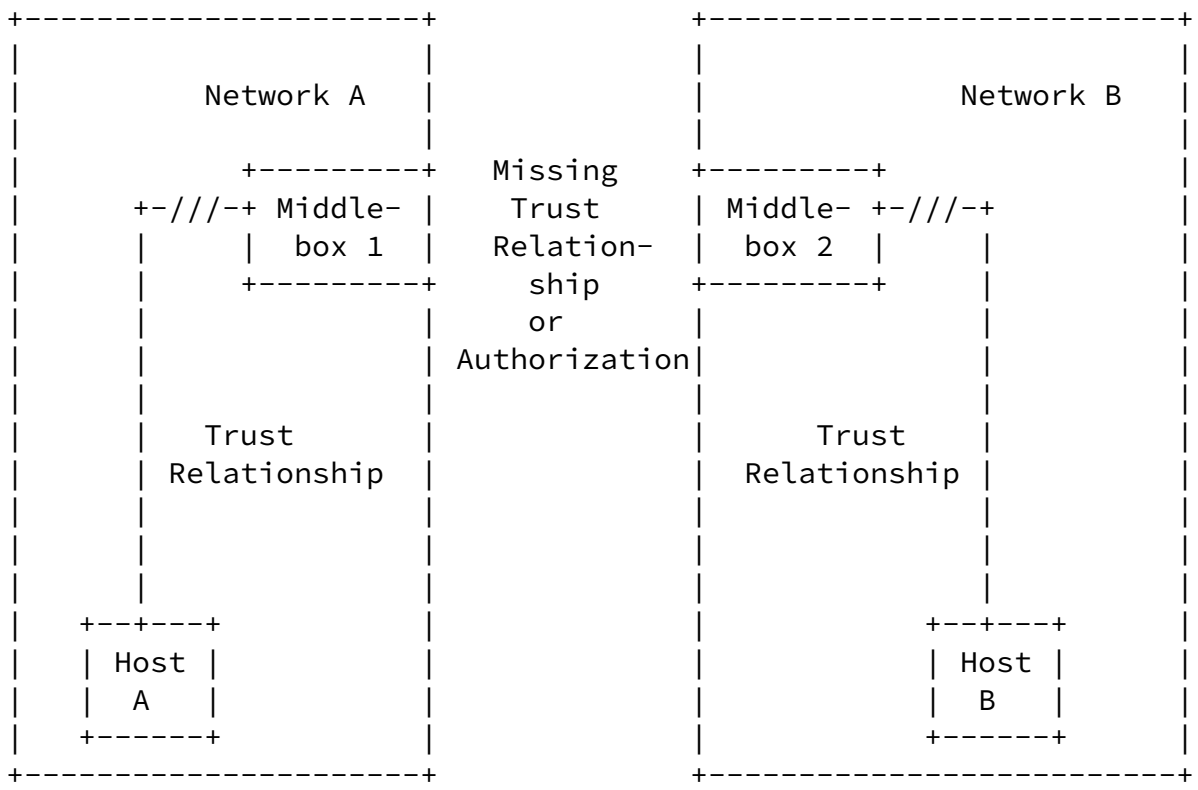


Figure 28: Missing Network-to-Network Trust Relationship

Figure 28 illustrates a problem whereby an external node is not allowed to manipulate (create, delete, query, etc.) packet filters at a Firewall. Opening pinholes is only allowed for internal nodes or with a certain authorization permission. Hence the solution alternatives in Section 3.2.2 focus on establishing the necessary trust with cooperation of internal nodes.

A.2 Relationship with routing

The data path is following the "normal" routes. The NAT/FW devices along the data path are those providing the service. In this case the service is something like "open a pinhole" or even more general "allow for connectivity between two communication partners". The benefit of using path-coupled signaling is that the NSIS NATFW NSLP does not need to determine what middleboxes or in what order the data flow will go through.

Creating NAT bindings modifies the path of data packets between two end points. Without NATs involved, packets flow from endhost to endhost following the path given by the routing. With NATs involved, this end-to-end flow is not directly possible, because of separated address realms. Thus, data packets flow towards the external IP

address at a NAT (external IP address may be a public IP address).

Other NSIS NSLPs, for instance QoS NSLP, which do not interfere with routing - instead they only follow the path of the data packets.

A.3 Affected Parts of the Network

NATs and Firewalls are usually located at the edge of the network, whereby other signaling applications affect all nodes along the path. One typical example is QoS signaling where all networks along the path must provide QoS in order to achieve true end-to-end QoS. In the NAT/Firewall case, only some of the domains/nodes are affected (typically access networks), whereas most parts of the networks and nodes are unaffected (e.g. the core network).

This fact raises some questions. Should an NSIS NTLP node intercept every signaling message independently of the upper layer signaling application or should it be possible to make the discovery procedure more intelligent to skip nodes. These questions are also related to the question whether NSIS NAT/FW should be combined with other NSIS signaling applications.

A.4 NSIS backward compatibility with NSIS unaware NAT and Firewalls

Backward compatibility is a key for NSIS deployments, as such the NSIS protocol suite should be sufficiently robust to allow traversal of none NSIS aware routers (QoS gates, Firewalls, NATs, etc).

NSIS NATFW NSLP's backward compatibility issues are different than the NSIS QoS NSLP backward compatibility issues, where an NSIS unaware QoS gate will simply forward the QoS NSLP message. An NSIS unaware Firewall rejects NSIS messages, since Firewalls typically implement the policy "default to deny".

The NSIS backward compatibility support on none NSIS aware Firewall would typically consist of configuring a static policy rule that allows the forwarding of the NSIS protocol messages (either protocol type if raw transport mode is used or transport port number in case a transport protocol is used).

For NATs backward compatibility is more problematic since signaling messages are forwarded (at least in one direction), but with a

changed IP address and changed port numbers. The content of the NSIS signaling message is, however, unchanged. This can lead to unexpected results, both due to embedded unchanged local scoped addresses and none NSIS aware Firewalls configured with specific policy rules allowing forwarding of the NSIS protocol (case of transport protocols are used for the NTLP). NSIS unaware NATs must be detected to maintain a well-known deterministic mode of operation for all the involved NSIS entities. Such a "legacy" NAT detection

procedure can be done during the NSIS discover procedure itself.

Based on experience it was discovered that routers unaware of the Router Alert IP option [RFC 2113] discarded packets, this is certainly a problem for NSIS signaling.

A.5 Authentication and Authorization

For both types of middleboxes, Firewall and NAT security is a strong requirement. Authentication and authorization means must be provided.

For NATFW signaling applications it is partially not possible to do authentication and authorization based on IP addresses. Since NATs change IP addresses, such an address based authentication and authorization scheme would fail.

A.6 Directional Properties

There two directional properties that need to be addressed by the NATFW NSLP:
- Directionality of the data
- Directionality of NSLP signaling

Both properties are relevant to NATFW NSLP aware NATs and Firewalls.

With regards to NSLP signaling directionality: As stated in the previous sections, the authentication and authorization of NSLP signaling messages received from hosts within the same trust domain (typically from hosts located within the security perimeter delimited by Firewalls) is normally simpler than received messages sent by hosts located in different trust domains.

The way NSIS signaling messages enters the NSIS entity of a Firewall (see Figure 2) might be important, because different policies might apply for authentication and admission control.

Hosts deployed within the secured network perimeter delimited by a Firewall, are protected from hosts deployed outside the secured network perimeter, hence by nature the Firewall has more restrictions on flows triggered from hosts deployed outside the security perimeter.

A.7 Addressing

A more general problem of NATs is the addressing of the end-point. NSIS signaling message have to be addressed to the other end host to follow data packets subsequently sent. Therefore, a public IP

address of the receiver has to be known prior to sending an NSIS message. When NSIS signaling messages contain IP addresses of the sender and the receiver in the signaling message payloads, then an NSIS entity must modify them. This is one of the cases, where a NSIS aware NATs is also helpful for other types of signaling applications e.g. QoS signaling.

A.8 NTLP/NSLP NAT Support

It must be possible for NSIS NATs along the path to change NTLP and/or NSLP message payloads, which carry IP address and port information. This functionality includes the support of providing mid-session and mid-path modification of these payloads. As a consequence these payloads must not be reordered, integrity protected and/or encrypted in a non peer-to-peer fashion (e.g. end-to-middle, end-to-end protection). Ideally these mutable payloads must be marked (e.g. a protected flag) to assist NATs in their effort of adjusting these payloads.

A.9 Combining Middlebox and QoS signaling

In many cases, middlebox and QoS signaling has to be combined at least logically. Hence, it was suggested to combine them into a single signaling message or to tie them together with the help of some sort of data connection identifier, later on referred as Session ID. This, however, has some disadvantages such as:

- NAT/FW NSLP signaling affects a much small number of NSIS nodes along the path (for example compared to the QoS signaling).

- NAT/FW signaling might show different signaling patterns (e.g. required end-to-middle communication).

- The refresh interval is likely to be different.

- The number of error cases increase as different signaling applications are combined into a single message. The combination of error cases has to be considered.

A.10 Inability to know the scenario

In Section 2.1 a number of different scenarios are presented. Data receiver and sender may be located behind zero, one, or more Firewalls and NATs. Depending on the scenario, different signaling approaches have to be taken. For instance, data receiver with no NAT and Firewall can receive any sort of data and signaling without any further action. Data receivers behind a NAT must first obtain a public IP address before any signaling can happen. The scenario

might even change over time with moving networks, ad-hoc networks or with mobility.

NSIS signaling must assume the worst case and cannot put responsibility to the user to know which scenario is currently applicable. As a result, it might be necessary to perform a "discovery" periodically such that the NSIS entity at the end host has enough information to decide which scenario is currently applicable. This additional messaging, which might not be necessary in all cases, requires additional performance, bandwidth and adds complexity. Additional, information by the user can provide information to assist this "discovery" process, but cannot replace it.

Appendix B. Acknowledgments

We would like to acknowledge Vishal Sankhla and Joao Girao for their input to this draft.

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in

this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

Full Copyright Statement

Acknowledgment

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