NSIS Working Group
Internet-Draft
Expires: January 17, 2005

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

July 19, 2004

NAT/Firewall NSIS Signaling Layer Protocol (NSLP)
draft-ietf-nsis-nslp-natfw-03

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 the Next Steps in Signaling (NSIS) working group.

Table of Contents

1. Introduction

Firewalls and Network Address Translators (NAT) have both been used throughout the Internet for many years, and they will remain present for the foreseeable future. Firewalls are used to protect networks against certain types of attacks from the outside, and in times of IPv4 address depletion, NATs virtually extend the IP address space. Both types of devices may be obstacles to many applications, since they only allow traffic created by a limited set of applications to traverse them (e.g., most HTTP traffic, and client/server applications), due to the rather static properties of those protocols. Other applications, such as IP telephony and most other peer-to-peer applications with more dynamic properties, create traffic which is unable to traverse NATs and Firewalls unassisted. In practice, the traffic from many applications cannot traverse Firewalls or NATs, even if they work autonomously in an attempt to restore the transparency of the network.

Several solutions to enable applications to traverse such entities have been proposed and are currently in use. Typically, application level gateways (ALG) have been integrated with the Firewall or NAT to

configure the Firewall or NAT dynamically. Another approach is middlebox communication (MIDCOM, currently under standardization at the IETF). In this approach, ALGs external to the Firewall or NAT configure the corresponding entity via the MIDCOM protocol [7]. Several other work-around solutions are available as well, such as STUN [35] and TURN [37]. However, all of these approaches introduce other problems that are hard to solve, such as dependencies on the type of NAT implementation (full-cone, symmetric, ...), or dependencies on a certain network topology.

NAT and Firewall (NATFW) signaling share a property with Quality of Service (QoS) signaling. Namely, both require that any device on the data path that is involved in QoS or NATFW treatment of data packets is reached. For both, NATFW and QoS, signaling travels path-coupled, meaning that the signaling messages follow exactly the same path that the data packets take. RSVP [14] is an example of a current QoS signaling protocol that is path-coupled.

This memo defines a path-coupled signaling protocol for NAT and Firewall configuration within the framework of NSIS, called the NATFW NSIS Signaling Layer Protocol (NSLP). The general requirements for NSIS are defined in [2]. The general framework of NSIS is outlined in [1]. It introduces the split between an NSIS transport layer and an NSIS signaling layer. The transport of NSLP messages is handled by an NSIS Network Transport Layer Protocol (NTLP, with GIMPS [3] being the implementation of the abstract NTLP). The signaling logic for QoS and NATFW signaling is implemented in the different NSLPs.

The QoS NSLP is defined in [4], while the NATFW NSLP is defined in this document.

The NATFW NSLP is designed to request 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. Every NATFW NSLP NAT/Firewall along the data path intercepts these messages, processes them, 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 CREATE mode in which state at firewalls and NATs is created. In the above example,

this takes place in the direction from the data sender to the data receiver. The other mode is the RESERVE mode. In this mode, NATs are discovered by the NSLP/NTLP signaling messages, and a publicly reachable IP address and a port number are reserved at each NAT. This mode enables hosts located in a private addressing realm delimited by a NAT to receive data originated in the public network. Both modes create NATFW NSLP and NTLP state in network entities. NTLP state allows signaling messages to travel in the forward (downstream) and the reverse (upstream) direction along the path between an NAT/Firewall NSLP sender and a corresponding receiver. NAT bindings and firewall rules are NAT/Firewall specific state. This state is managed using a soft-state mechanism, i.e., it expires unless it is refreshed every now and then by a certain message. If state is to be deleted explicitly before it automatically expires, another message can be used for that. To find out which state is currently installed in NSIS NAT/Firewall nodes, a query message can be used at any time.

Section 2 describes the network environment for NATFW NSLP signaling, highlighting the trust relationships and authorization required. Section 3 defines the NATFW signaling protocol. Section 5 defines the messages and and message components. In the remaining parts of the main body of the document, Section 6 covers transition issues, while Section 7 addresses security considerations, with more extensive discussions of security issues currently being contained in [20]. Currently unsolved problems and challenges are listed and discussed in Appendix A. Please note that readers familiar with Firewalls and NATs and their possible location within 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 a number of terms defined in [2]. The following additional terms are used:

- NSIS NAT Forwarding State: This term refers to a state used to forward the NSIS signaling message beyond the targeted destination address.

- Policy rule: 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" [38]. 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", "deny" or action "allow".
- Firewall: A packet filtering device that matches packets 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 between hosts (see [8]). Network Address Translators are devices that perform this method.
- Middlebox: "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" [12]. In the context of this document and in NSIS, the term middlebox 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 that makes a resource request, usually as a result of user application request.
- NSIS Responder (NR): The signaling entity that acts as the final destination for the signaling. It can optionally interact with applications as well.
- NSIS Forwarder (NF): A signaling entity between an NI and an NR which propagates NSIS signaling further through the network.
- Receiver (DR or R): The node in the network that is receiving the data packets of a flow.
- Sender (DS or S): The node in the network that is sending the data packets of a flow.
- NATFW NSLP session: An 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 consists of NSLP state and associated policy rules at a middlebox.
- NSIS peer or peer: An NSIS node with which an NSIS adjacency has been created as defined in [3].

- Edge NAT: An edge NAT is a NAT device that is reachable from the public Internet and that has a globally routable IP address.
- Edge Firewall: An edge Firewall is a Firewall device that is located on the demarcation line of an administrative domain.
- Public Network: "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" [8].
- Private/Local Network: "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" [8]. IP address space allocation for private networks is recommended in [36]
- Public/Global IP address: An IP address located in the public network according to Section 2.7 of [8].
- Private/Local IP address: An IP address located in the private network according to Section 2.8 of [8].
- Initial CREATE: A CREATE message creating a new session.

1.2 Middleboxes

The term middlebox covers a range of devices which intercept the flow of packets between end hosts and perform actions other than standard forwarding expected in an IP router. As such, middleboxes fall into a number of categories with a wide range of functionality not all of which is pertinent to the NATFW NSLP. Middlebox categories in the scope of this memo are Firewalls that filter data packets against a set of filter rules, and NATs that translate packet addresses from one address realm to another address realm. Other categories of middleboxes, such as 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 these types of NATs:
- traditional NAT (basic NAT and NAPT)
- Bi-directional NAT
- Twice-NAT
- Multihomed NAT

For definitions and a detailed discussion about the characteristics of each NAT type please see [8].

Both types of middleboxes under consideration here use policy rules to make a decision on data packet treatment. Policy rules consist of a flow identifier (which is typically a 5-tuple) and an associated action; data packets matching the flow identifier are subjected to the policy rule action. A 5-tuple selector matches the following

fields of a packet to configured values:
- Source and destination IP addresses
- Transport protocol number
- Transport source and destination port numbers

For further examples of flow identifiers see Section 5.1 of [3].

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

Actions for NATs include (amongst many others):
- Change source IP address and transport port number to a globally routeable IP address and associated port number.
- Change destination IP address and transport port number to a private IP address and associated port number.

1.3 Non-Goals

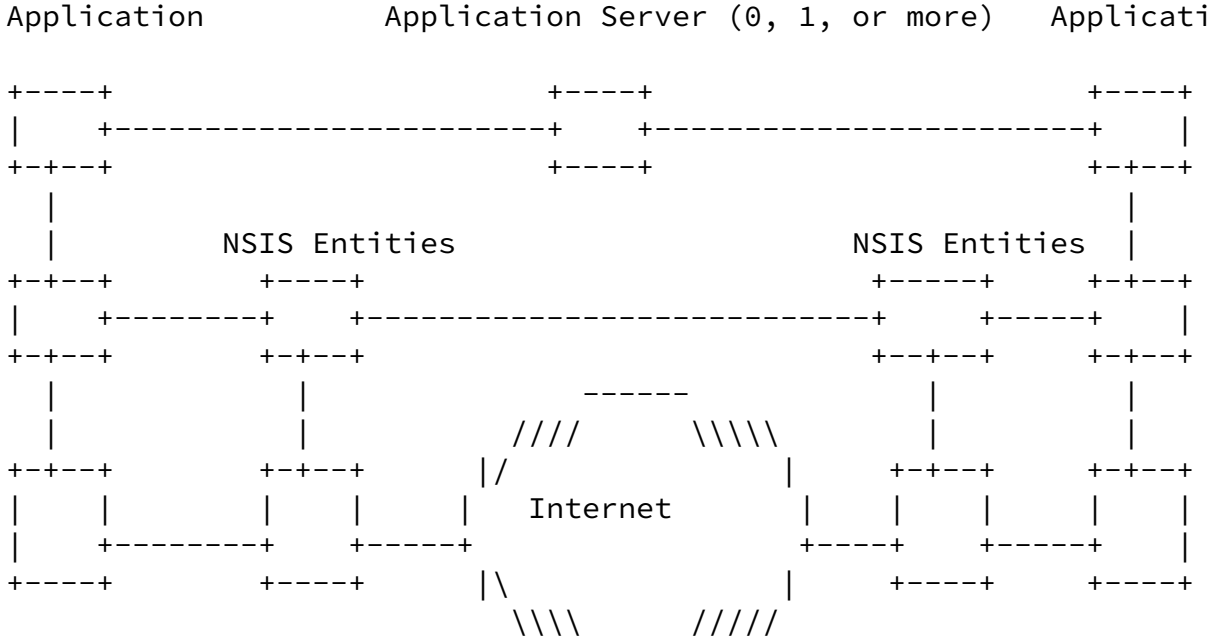
Traversal of non-NSIS and non-NATFW NSLP aware NATs and Firewalls is outside the scope of this document.
Only Firewalls and NATs are considered in this document, any other types of devices, for instance IPSec security gateway, are out of scope.
The exact implementation of policy rules and their mapping 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 categorized 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 to as security gateways) since per-flow signaling is rather uncommon in this environment.
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 outside the scope of this document.
In mobility scenarios, a common problem is the traversal of a security gateway at the edge of a corporate network. Network administrators allow only authenticated data to enter the network. A problem statement for the traversal of these security gateways in the context of Mobile IP can be found in [28]). This topic is not within the scope of the present document.

1.4 General Scenario for NATFW Traversal

The purpose of NSIS NATFW signaling is to enable communication between endpoints across networks even in the presence of NAT and

Firewall middleboxes. It is assumed that these middleboxes will be statically configured in such a way that NSIS NATFW signaling messages themselves are allowed to traverse them. NSIS NATFW NSLP signaling is used to dynamically install additional policy rules in all NATFW 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 end hosts are located behind middleboxes (NAT/FW in Figure 1). Applications located at these end hosts try to establish communication with corresponding applications on other such end hosts. They trigger the NSIS entity at the local host to provide for middlebox traversal along the prospective data path (e.g., via an API call). The NSIS entity in turn uses NSIS NATFW NSLP signaling to establish policy rules along the data path, allowing the data to travel from the sender to the receiver unobstructed.



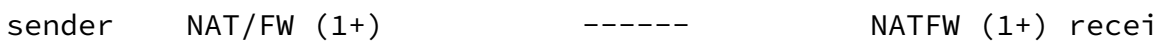


Figure 1: Generic View on NSIS in a NAT / Firewall case

For end-to-end NATFW signaling, it is necessary that each firewall and each NAT along the path between the data sender and the data receiver implement the NSIS NATFW NSLP. There might be several NATs and FWs in various possible combinations on a path between two hosts. Section 2 presents a number of likely scenarios with different combinations of NATs and firewalls.

2. Network Deployment Scenarios using NATFW NSLP

This section introduces several scenarios for middlebox placement within IP networks. Middleboxes are typically found at various different locations, including at Enterprise network borders, within enterprise networks, as mobile phone network gateways, etc. Usually, middleboxes are placed rather towards the edge of networks than in network cores. Firewalls and NATs may be found at these locations either alone, or they may be combined; other categories of middleboxes may also be found at such locations, possibly combined with the NATs and/or Firewalls. To reduce the number of network elements needed, combined Firewall and NATs have been made available.

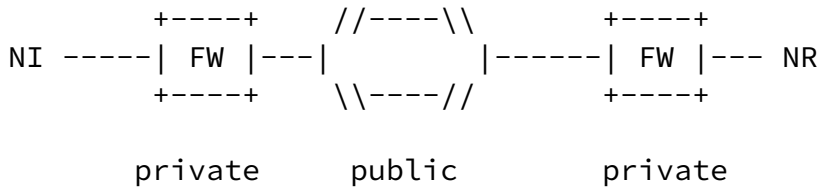
NSIS initiators (NI) send NSIS NATFW NSLP signaling messages via the regular data path to the NSIS responder (NR). On the data path, NATFW NSLP signaling messages reach different NSIS peers that implement the NATFW NSLP. Each NATFW NSLP node processes the signaling messages according to Section 3 and,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

Each of the following sub-sections 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 Firewall Traversal

This section describes a scenario with Firewalls only; NATs are not involved. Each end host is behind a Firewall. The Firewalls are

connected via the public Internet. Figure 2 shows the topology. The part labeled "public" is the Internet connecting both Firewalls.



FW: Firewall
NI: NSIS Initiator
NR: NSIS Responder

Figure 2: Firewall Traversal Scenario

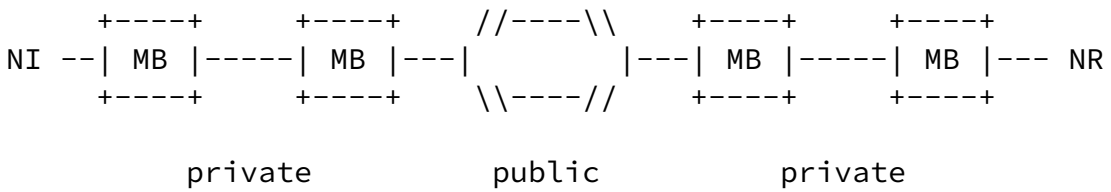
Each Firewall on the data path must provide traversal service for NATFW NSLP in order to permit the NSIS message to reach the other end host. All Firewalls process NSIS signaling and establish appropriate

policy rules, so that the required data packet flow can traverse them.

2.2 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 generically labeled as MB (for middlebox), since those devices definitely implement NAT functionality, but can implement firewall functionality as well.

Only two middleboxes MB are shown in Figure 3 at each side, but in general, any number of MBs on each side must be considered.



MB: Middlebox

NI: NSIS Initiator
NR: NSIS Responder

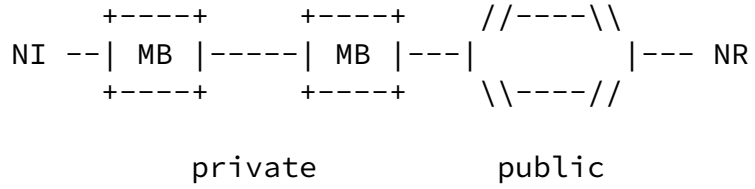
Figure 3: NAT with two Private Networks Scenario

Signaling traffic from NI to NR has to traverse all the middleboxes on the path, and all the middleboxes must be configured properly to allow NSIS signaling to traverse them. The NATFW signaling must configure all middleboxes and consider any address translation that will result from this configuration in further signaling. The sender (NI) has to know the IP address of the receiver (NR) in advance, otherwise it will not be possible to send any NSIS signaling messages 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 previously installed on the NAT that subsequently allows packets reaching 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.

2.3 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 generic MB, see discussion in Section 2.2). The receiver is located in the public Internet.



MB: Middlebox
NI: NSIS Initiator
NR: NSIS Responder

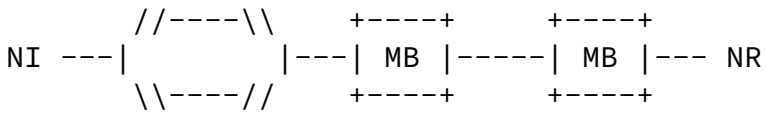
Figure 4: NAT with Private Network on Sender Side Scenario

The traffic from NI to NR has to traverse middleboxes only 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 3.3.2 NSIS can also provide help for this procedure.

2.4 NAT with Private Network on Receiver Side Scenario

The application instance receiving data is behind one or more NATs shown as MB (see discussion in Section 2.2).



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 middleboxes in the receiver side private network.

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

2.5 Both End Hosts behind twice-NATs

This is a special case, where the main problem consists of detecting that both end hosts are logically within the same address space, but are also in two partitions of the address realm on either side of a twice-NAT (see [8] for a discussion of twice-NAT functionality).

Sender and receiver are both within a single private address realm but the two partitions potentially have overlapping IP address ranges. Figure 6 shows the arrangement of NATs. This is a common configuration in networks, particularly after the merging of companies that have used the same private address space, resulting in overlapping address ranges.

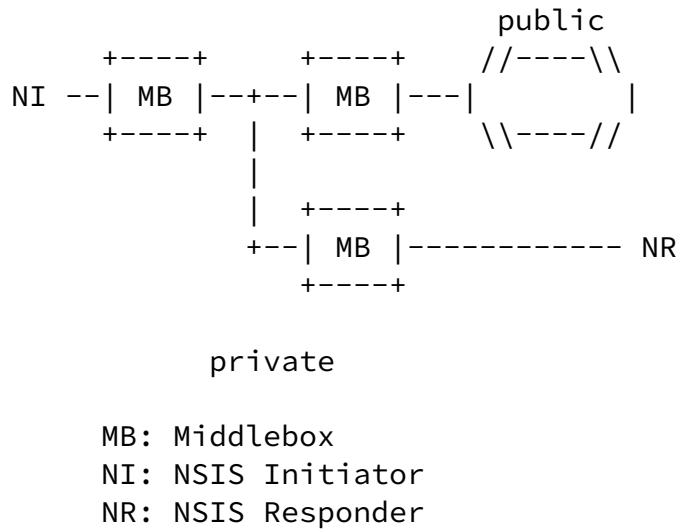


Figure 6: NAT to Public, Sender and Receiver on either side of a twice-NAT Scenario

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.

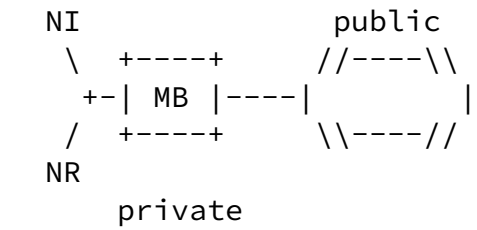
This scenario requires assistance of application level entities, such as a DNS server. The application level gateways must handle requests that are based on symbolic names, and configure the middleboxes so that data packets are correctly forwarded from NI to NR. The configuration of those middleboxes may require other middlebox

communication protocols, like MIDCOM [7]. NSIS signaling is not

required in the twice-NAT only case, since the middleboxes of the twice-NAT type are normally 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 Firewalls on the path. The NSIS signaling protocol must be at least robust enough to survive this scenario.

2.6 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 2.4 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 twice: 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 in which both sender and receiver obtain a public IP address at the NAT, and the communication path is certainly not optimal in this case.



MB: Middlebox
NI: NSIS Initiator
NR: NSIS Responder

Figure 7: NAT to Public, Both Hosts Behind Same NAT

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

2.7 IPv4/v6 NAT with two Private Networks

This scenario combines the use case described in Section 2.2 with the IPv4 to IPv6 transition scenario involving address and protocol translation, i.e., using Network Address and Protocol Translators

(NAT-PT, [11]).

The difference from the other scenarios is the use of IPv6 to IPv4 (and vice versa) address and protocol translation. Additionally, the base NTLP must support transport of messages in mixed IPv4 and IPv6 networks where some NSIS peers provide translation.

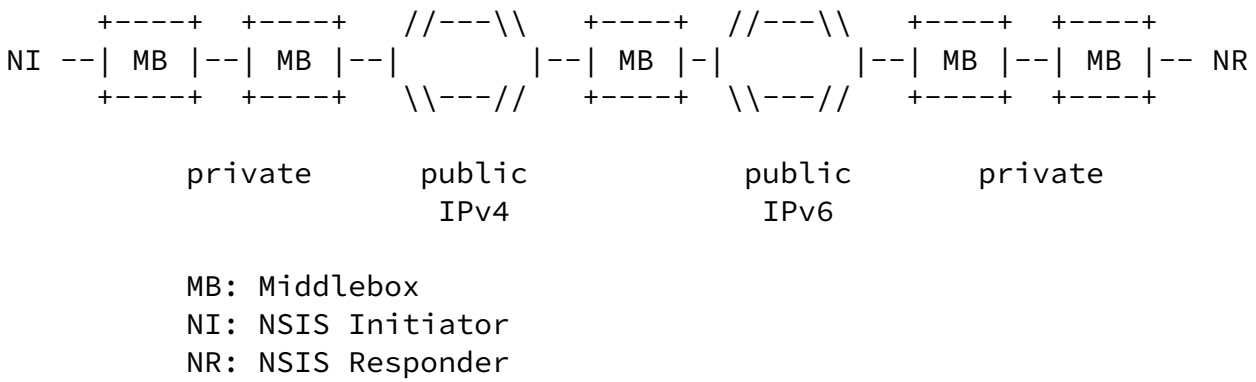


Figure 8: IPv4/v6 NAT with two Private Networks

This scenario needs the same type of application level support as described in Section 2.5, and so the issues relating to twice-NATs apply here as well.

2.8 Multihomed Network with NAT

The previous sub-sections sketched network topologies where several NATs and/or Firewalls are ordered sequentially on the path. This section describes a multihomed scenario with two NATs placed on alternative paths to the public network.

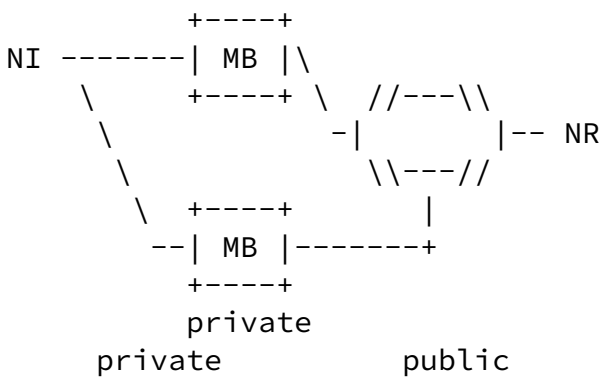


Figure 9: Multihomed Network with Two NATs

Depending on the destination or load balancing requirements, either one or the other middlebox is used for the data flow. Which middlebox is used depends on local policy or routing decisions. NATFW NSLP must be able to handle this situation properly, see Section 3.3.2 for an expanded discussion of this topic with respect to NATs.

3. Protocol Description

This section defines messages, objects, and protocol semantics for the NATFW NSLP. Section 3.1 introduces the base constituent element of a session state, the policy rule. Section 3.2 introduces the protocol and the protocol behavior is defined in Section 3.3. Section 5 defines the syntax of the messages and objects.

3.1 Policy Rules

Policy rules, bounded to a session, 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. For NATs the policy rule consists of action 'translate this address to realms address pool' 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 allow or deny action, and additional information. The filter specification is carried within NTLP's message routing information (MRI) and additional information is carried in NSLP's objects. Additional information is, for example, the lifetime of a policy rule or session.

3.2 Basic protocol overview

The NSIS NATFW NSLP 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. NSIS forwarders that do not have any NATFW NSLP functions just forward these packets when they have no interest (which is expected to happen in most cases).

A Data Sender (DS), intending to send data to a Data Receiver (DR) must first start its NATFW NSLP signaling. In the next step, the NI at the data sender (DS) starts NSLP signaling towards the address of data receiver DR (see Figure 10). Although the above NATFW NSLP usage is expected to be the most common, this specification does not prevent scenarios where the data sender and NI reside on different hosts.

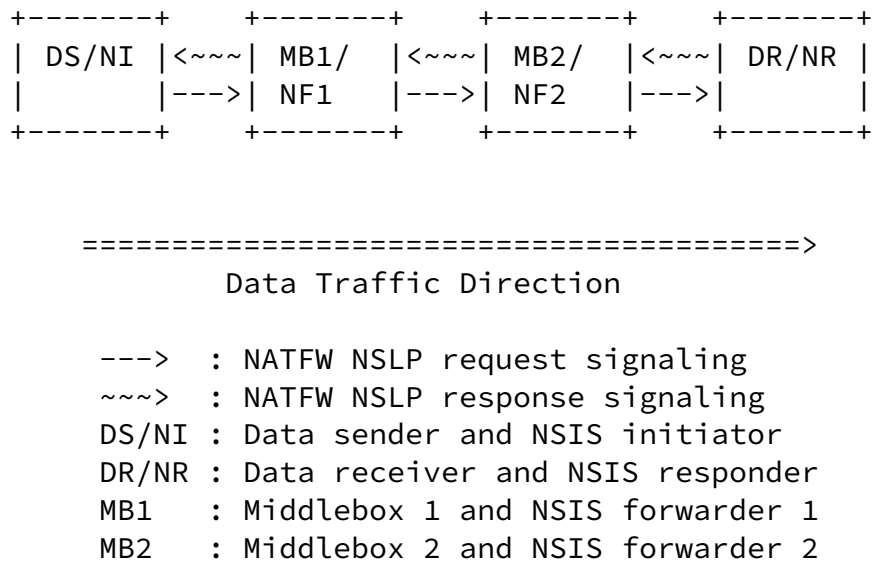


Figure 10: General NSIS signaling

The sequence of NSLP events is as follows:

o NSLP request messages are processed each time a NF with NATFW NSLP support is passed. These 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 them if applicable. NSIS responders generate response messages and send them hop-by-hop back to the NI via the same chain of NFs (traversal of the same NF chain is guaranteed through the established reverse message routing state in the NTLP).
- The response message is processed at each NF implementing NATFW NSLP.
- Once the NI has received a successful response, the Data Sender can start sending its data flow to the Data Receiver.

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 the NR and the NI are located in different address realms and the NR is behind a NAT, the NI cannot signal to the NR directly. The NR is not reachable from the NIs and thus no NATFW signaling messages can be sent to the DR's address. Therefore, the NR must first obtain a NAT binding that is reachable for the NI. Once the NR has determined a public IP address, it forwards this information to the

DS via a separate protocol (such as SIP). This application layer signaling, out of scope of the NATFW NSLP, may involve third parties that assist in exchanging these messages.

NATFW NSLP signaling supports this scenario by using the RESERVE mode of operation :

1. The NR determines a public address by signaling on the reverse path (NR towards NI) 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.
2. The NI signals directly the NR as NI would do if there is no NAT in between, and creates policy rules at middleboxes. Note, that the reservation mode will make reservations only, which will be "activated" by the signaling from NI towards NR. The first mode is detailed in the Section 3.3.2

The protocol works on a soft-state basis, meaning 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 NATFW nodes involved will have to specifically request a session extension. An explicit NATFW NSLP state deletion capability is also provided by the protocol.

Middleboxes should return an error in case of a failure, such that appropriate actions can be taken; this ability would allow debugging and error recovery. Error messages could be sent upstream (for errors related to received messages as well as asynchronous error notification messages) towards the NI as well as downstream towards the NR (case of asynchronous error notification messages).

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

3.3 Protocol Operations

This section defines the protocol operations, how to create sessions, maintain them, and how to reserve addresses. All the protocols messages require C-mode handling by the NTLP and cannot be piggybacked to D-mode NTLP messages used during the NTLP path discovery/refresh phase. The protocol messages NTLP usage is described in more details within Section 5.

The protocol uses six messages:

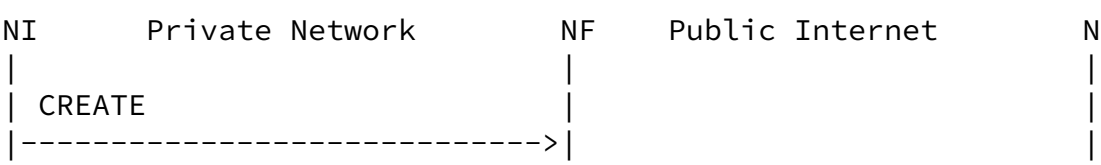
- CREATE: a request message used for creating, changing, refreshing and deleting NATFW NSLP sessions.

- RESERVE-EXTERNAL-ADDRESS (REA): a request message used for reserving an external address
- RESPONSE: used to response to CREATE, REA and QUERY messages with Success or Error information
- QUERY: a request message used by authorized NATFW NEs for querying NATFW on installed stated
- NOTIFY: an asynchronous message used by NATFW NEs to alert upstream and/or downstream NATFW NEs about specific events (mainly failures).
- TRIGGER: a message sent upstream to trigger CREATE messages to be sent.

The following sections will present the semantics of these messages by exhibiting their impact on the protocol state machine.

3.3.1 Creating Sessions

Allowing two hosts to exchange data even in the presence of middleboxes is realized in the NATFW NSLP by the 'CREATE ' request message. The data sender generates a CREATE message as defined in Section 5.4.1 and hands it to the NTLP. The NTLP forwards the whole message on the basis of the message routing information towards the NR. Each NSIS forwarder along the path that is implementing NATFW NSLP, processes the NSLP message. Forwarding is thus managed NSLP hop-by-hop but may pass transparently through NSIS forwarders which do not contain NATFW NSLP functionality and non-NSIS aware routers between NSLP hop waypoints. When the message reaches the NR, the NR can accept the request or reject it. NR generates a response to the request and this response is transported hop-by-hop towards the NI. NATFW NSLP forwarders may reject requests at any time. Figure 11 sketches the message flow between NI (DS), a NF (NAT), and NR (DR).



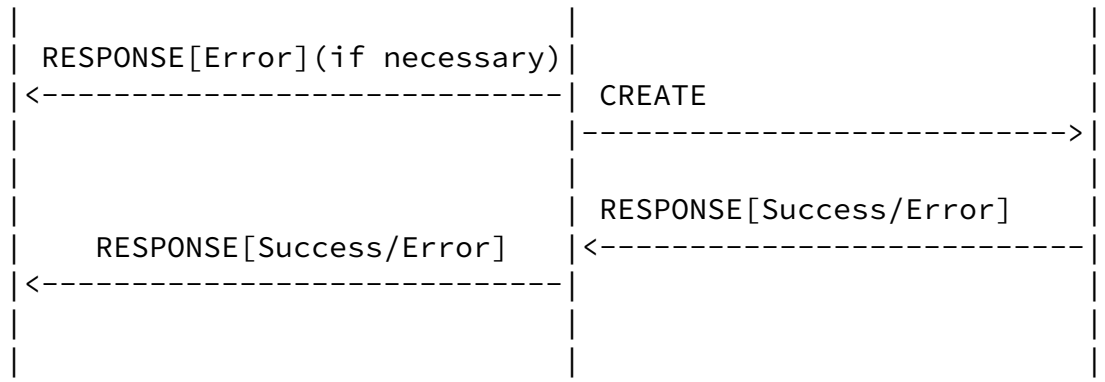


Figure 11: Creation message flow

Since the CREATE message is used for several purposes within the lifetime of a session, there are several processing rules for NATFW NEs when generating and receiving CREATE messages. The different processing methods depend not only if the CREATE is used to create, modify, refresh or delete a session but also on the node at which the processing happens. For an initial CREATE message the processing of CREATE messages is different for every NSIS node type:

- NSLP initiator: NI only generates initial CREATE messages and hands them over to the NTLP. After receiving a successful response, the data path is configured and the DS can start sending its data to the DR. After receiving an 'error' response message the NI MAY try to generate the CREATE message again or give up, depending on the error condition.
- NATFW NSLP forwarder: NFs receiving an initial CREATE message MUST first check authentication and authorization before any further processing is executed. The NF SHOULD check with its local policies if it can accept the desired policy rule given the combination of the NTLP's 'Message-Routing-Information' (MRI) [3] (the flow description information) and the CREATE payload (behavior to be enforced on the packet stream). The NSLP message processing depends on the middlebox type:
  - NAT: When the initial CREATE message is received at the public side of the NAT, it looks for a reservation made in advance, by using a REA message Section 3.3.2 , that matches the destination address/port of the MRI provided by the NTLP. If no reservation had been made in advance the NSLP SHOULD return an error response 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 (private in most cases) address of the next NSIS node. When the initial CREATE message, for a new session, 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 from the newly reserved binding.

* Firewall: When the initial CREATE 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. There is a difference between requests from trusted (authorized NIs) and un-trusted (un-authorized NIs); requests from trusted NIs will be pre-authorized, whereas requests from un-trusted NIs will not be pre-authorized. This difference is required to speed-up the protocol operations as well as for the proxy mode usage (please refer to Section 3.4 and [17]).
* 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 into account the order of packet processing in the Firewall and NAT functions within the device. This will be referred to as 'order of functions' and is generally different depending on whether the packet arrives at the external or internal side of the middlebox.

o NSLP receiver: NRs receiving initial CREATE messages MUST reply with a 'success' (response object has success information) RESPONSE message if they accept the CREATE request message. Otherwise they SHOULD generate a RESPONSE message with an error code. RESPONSE messages are sent back NSLP hop-by-hop towards the NI, independently of the response codes, either success or error.

Policy rules at middleboxes MUST be only installed upon receiving a successful response. This is a countermeasure to several problems, for example wastage of resources due to loading policy rules at intermediate NF when the CREATE message does not reach the final the NR for some reason.

3.3.2 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 as described in Section 3.3.1. 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) the problem 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.

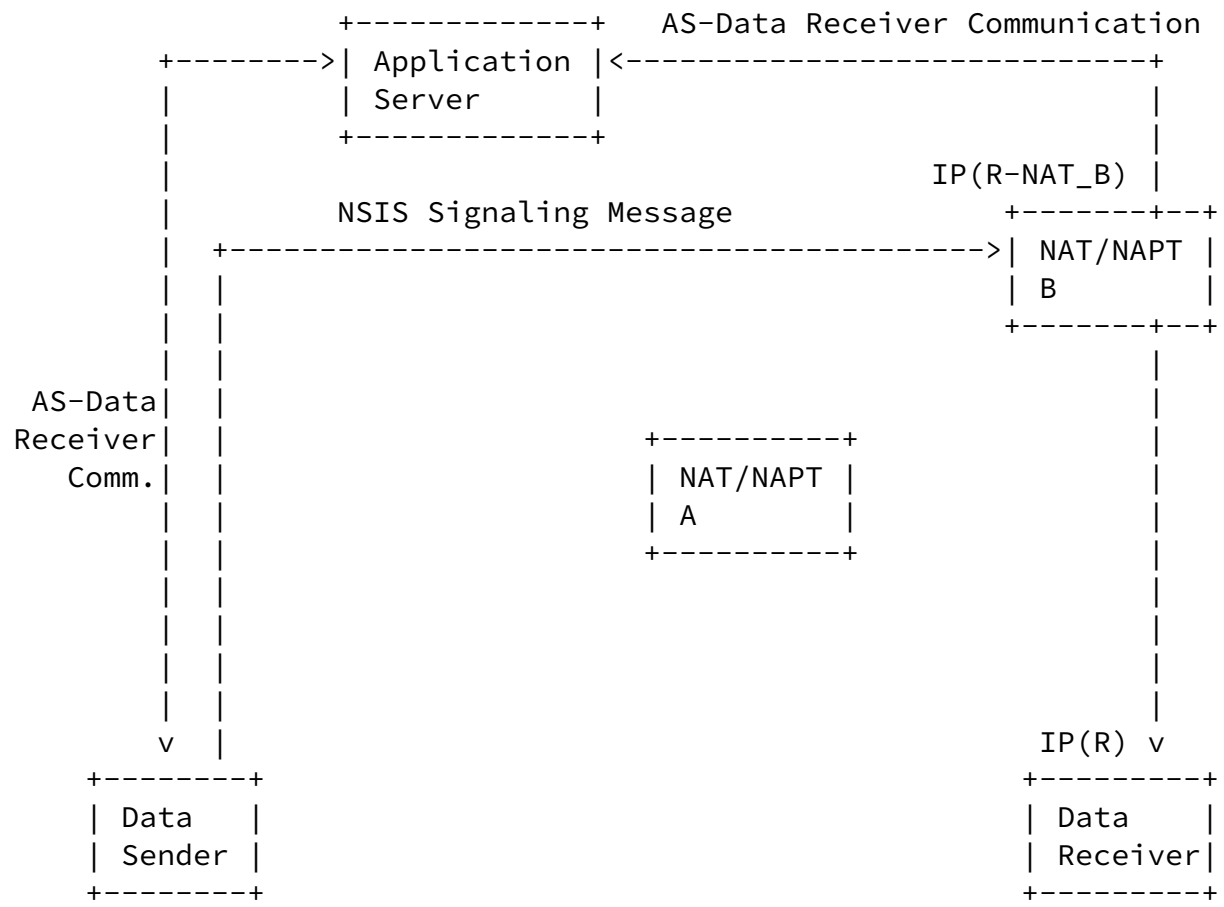


Figure 12: The Data Receiver behind NAT problem

Figure 12 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 between the application server each of the two end points (data sender and data receiver) enables the two end hosts to learn each others IP address. 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.

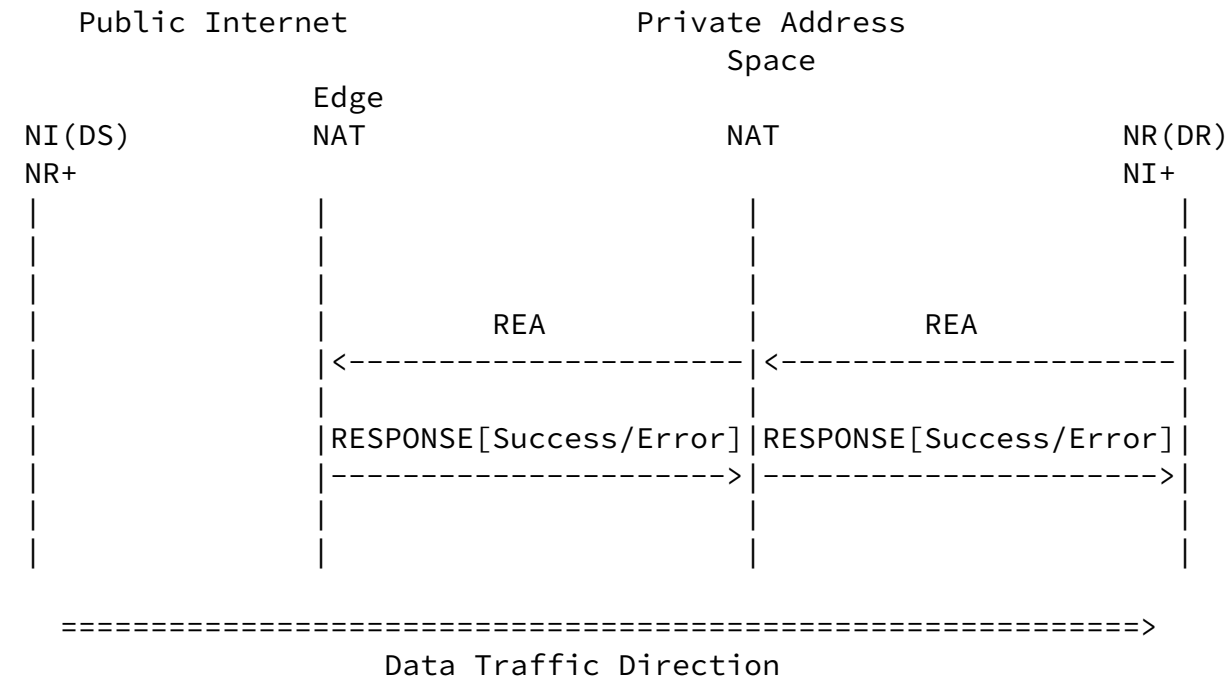


Figure 13: Reservation message flow

Figure 13 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 data receiver (DR) for the anticipated data traffic is the NSIS initiator (NI+) for the RESERVE-EXTERNAL-ADDRESS (REA) message, but becomes the NSIS responder (NR) for following CREATE messages.
- The actual data sender (DS) will be the NSIS initiator (NI) for later CREATE messages and may be the NSIS target of the signaling (NR+).
- The actual target of the REA message may be an arbitrary address, the Opportunistic Address (OA) that would force the message to get intercepted by the far outmost NAT in the network. .

The NI+ agent (could be on the data receiver DR or on any other host

within he private network) sends a the REA message targeted to the Opportunistic Address (OA). The OA selection for this message is discussed in Section 3.8. The message routing for the REA message is in the reverse direction to the normal message routing used for path-coupled signaling where the signaling is sent downstream (as opposed to upstream in this case). When 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 REA signaling message creates NSIS NAT Forwarding State at any intermediate NSIS NAT node(s) encountered. 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 is located at the outer perimeter of the private network. Forwarding of the REA ' message beyond this entity is not necessary, and should be prohibited as it provides information on the capabilities of internal hosts.

The edge NAT device responds to the REA message with a RESPONSE message containing a success object carrying the public reachable IP address/port number.

Processing of REA messages is specific to the NSIS node type:

- NSLP initiator: NI+ only generate REA messages and should never

receive them.

- NSLP forwarder: NSLP forwarders receiving REA messages MUST first check authentication and authorization before any further processing is executed. The NF SHOULD check with its local policies if it can accept the desired policy rule given by NTLP's message routing information (MRI). Further processing depends on the middlebox type:
  - NAT: NATs check whether the message is received at the external (public in most cases) address or at the internal (private) address. If received at the internal address a NF MAY generate a RESPONSE message with an error of type 'REA received from outside'. If received at the internal 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 message with the external address and port information is generated. If it is not an edge-NAT, the NSLP message is forwarded further with the translated IP address/port (if required by the NI+).
  - Firewall: Firewalls MUST not change their configuration upon a REA message. They simply MUST forward the message and MUST keep NTLP state. Firewalls that are configured as edge-Firewalls MAY return an error of type 'no NAT here'.

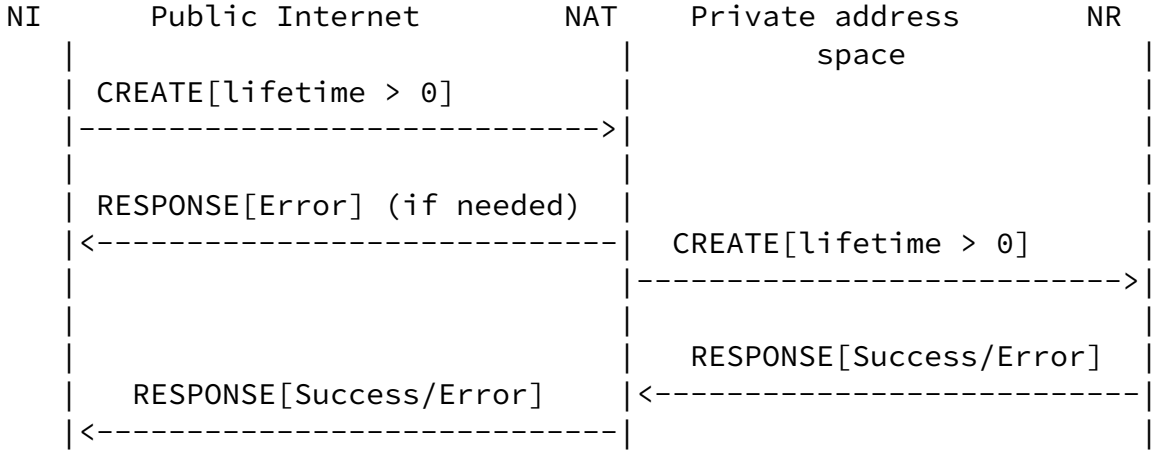
  - Combined NAT and Firewall: Processing at combined Firewall and NAT middleboxes is the same as in the NAT case.
- NSLP receiver: This type of message should never be received by any NR and it SHOULD be discarded silently.

Processing of a RESPONSE message with an external address object is different for every NSIS node type:

- NSLP initiator: Upon receiving a RESPONSE message with an external address object, the NI+ can use the IP address and port pairs carried 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 an NR and it SHOULD be discarded silently.
- Edge-NATs: This type of message should never be received by any Edge-NAT and it SHOULD be discarded silently.

3.3.3 NATFW Session refresh

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. The protocol uses a CREATE message to refresh sessions. Even if used for refresh purposes the CREATE message requires to be responded back, to allow the intermediate NFs to propose a refresh period that would align to their local policies. The NI sends CREATE messages destined for the NR. Upon reception by each NSIS forwarder, the state for the given session ID is extended by the session refresh period, a period of time calculated based on a proposed refresh message period. Extending lifetime of a session is calculated as current local time plus proposed lifetime value (session refresh period). Section 3.5 defines the process of calculating lifetimes in detail.



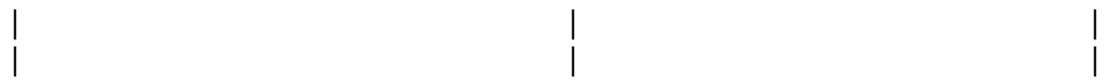


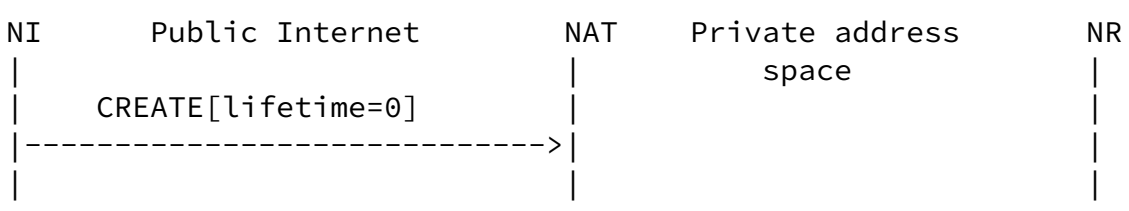
Figure 14: State Refresh Message Flow

Processing of session refresh CREATE messages is different for every NSIS node type:

- NSLP initiator: NI can generate session refresh CREATE messages before the session times out. The rate at which the refresh CREATE messages are sent and their relation to the session state lifetime are further discussed in Section 3.5. The message routing information and the extended flow information object MUST be set equal to the values of the initial CREATE request message.
- NSLP forwarder: NSLP forwarders receiving session refresh messages MUST first check authentication and authorization before any further processing is executed. The NF SHOULD check with its local policies if it 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: NRs accepting this session refresh CREATE message generate a RESPONSE message with response object set to success.

3.3.4 Deleting Sessions

NATFW NSLP sessions may be deleted at any time. NSLP initiators can trigger this deletion by using a CREATE messages with a lifetime value set to 0, as shown in Figure 15.



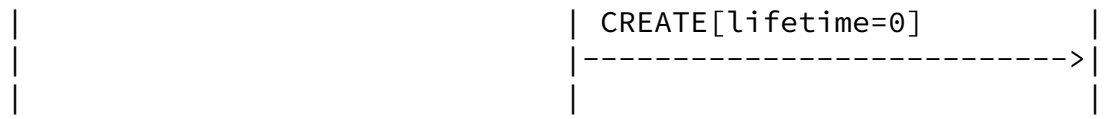


Figure 15: Delete message flow

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 CREATE request message with a lifetime value of 0, does not generate any response, neither positive nor negative, since there is no NSIS state left at the nodes along the path.

3.3.5 Reporting Asynchronous Events

NATFW NSLP forwarders and NATFW NSLP responders must have the ability to report asynchronous events to other NATFW NSLP nodes, especially reporting back to the NATFW NSLP initiator. Such asynchronous events may be premature session termination, changes in local polices, or any other reason that indicates change of the NATFW NSLP session state. Currently, only asynchronous session termination is defined as event, but other events may be defined in later versions of this memo.

NFs and NRs may generate NOTIFY messages upon asynchronous events, with a response object indicating the reason of the event. There are two suggested mode of operations:

1. NOTIFY messages are sent hop-by-hop upstream towards NI. Those NOTIFY messages may be sent downstream towards NR, if generated by a NF, if needed. TBD: Should there be a way to configure whether NOTIFY messages are sent downstream, too?
2. During session creation, via CREATE or REA, NIs may insert a special 'notify address' object into the NSLP message, indicating a node's address that should be notified about this event. TBD: When this object is used, is it desired to send the NOTIFY to both, NI and the other node? Sending to both could end up in one asynchronous event generating three messages: NOTIFY to NI (upstream), NOTIFY to NR (downstream), and NOTIFY to notify address.

Processing is different for every NATFW NSLP node type and only defined for asynchronous session termination events:

- NSLP initiator: NIs receiving NOTIFY messages MUST first check for authentication and authorization. After successfully doing so, NIs MUST remove the NSLP session as indicated by the NOTIFY message. NIs MUST NOT generate NOTIFY messages.
- NSLP forwarder: NFs receiving NOTIFY messages MUST first check for authentication and authorization. After successfully doing so, NFs MUST remove the NSLP session and corresponding policy rules immediately and MUST forward the NOTIFY message. NFs occurring an asynchronous event generate NOTIFY messages and set the response object to 'session termination' code. NOTIFY messages are sent hop-by-hop upstream towards NI (This depends on above mentioned design choice).
- NSLP responder: NRs may generate NOTIFY messages. NRs receiving NOTIFY messages MUST first check for authentication and authorization. After successfully doing so, NRs MUST remove the NSLP session immediately. NRs occurring an asynchronous event generate NOTIFY messages and set the response object to 'session termination' code. NOTIFY messages are sent hop-by-hop upstream towards NI (This depends on above mentioned design choice).

3.3.6 QUERY capabilities within the NATFW NSLP protocol

The NATFW NSLP provides query capabilities that could be used by:

- A session owner to track the session state, this would be used for diagnosis when no data packets were received and the policy rule was supposed to be created on the NATFW NFs.
- A superuser to track user activities, detect misbehaving users and blocking them from using the NATFW NSLP on the NATFW NFs within the network. When doing so it is recommended that the QUERY message be scoped to the limits of the administrative domain.

The QUERY message could be used to query the following information:

- Session information: session id, flow source, destination and status of the state listed in best status to worst status: up, high traffic (used to detect DOS attack or unexpected traffic rate), pending, down. The status of the policy rule indicate sufficient diagnosis information, in case more diagnosis information is required it could be provided by the NATFW NF logs. Session status is only provided by an NF if no session status was provided in the QUERY message or the NF's session status is worst than the one provided by the queried upstream NEs. The Session information could be retrieved by sending a QUERY against a specific session id, a flow source and destination or user identifier with session id or flow source and destination.
- User identifiers: the query would be used by a super-user to track activities of a suspected user, the query would return all the suspected user active sessions

QUERY message processing is different for every NATFW NSLP node type:

- NSLP initiator: NIs only generate QUERY messages, but never with session status information, in case received QUERY messages MUST be discarded.
- NSLP forwarder: NFs receiving QUERY messages MUST first check for authentication and authorization. After successfully doing so, NFs will behave differently depending on the QUERY.
  - if the QUERY is about a specific session: if it contains a session status the NF compares it to the current local session status; if no session status is provided in the QUERY message the NF will insert its own session status in the QUERY message. If the current local session status is worst, it will incorporate its own session status field in the QUERY message. Every NF will provide the flow description in case it was not inside the QUERY.
  - if the QUERY is about a specific user, the NF will gather all the user's sessions and provide a list of them.

  Once the message processing is done, if the message was not scoped then NF will forward the QUERY message to the next downstream node.
- NSLP responder: NRs (any node being the destination of the message)receiving QUERY messages MUST first check for authentication and authorization. After successfully doing so, NRs must process the message as the NFs and respond with a RESPONSE message to the NI. The RESPONSE message will travel along the established reverse path Message Routing State.

Responses to QUERY messages are processed differently for every NATFW NSLP node type:

- NSLP initiator: NIs receiving RESPONSEs to QUERY messages MUST first check for authentication and authorization. After successfully doing so, the objects within the RESPONSE messages are provided up to the application layers and the session state remains as it was unless the application triggers NATFW NSLP state changes.
- NSLP forwarder: NFs receiving RESPONSEs to QUERY messages MUST first check for authentication and authorization. After successfully doing so, NFs forward the message upstream without any interpretation.
- NSLP responder: if an NR received a RESPONSE to QUERY message it MUST discard it.

3.3.7 QUERY Message semantics

From a semantics perspective, the QUERY messages may require the following information incorporated within the messages:

- Session ID

- User ID
- Flow source (address and port) and destination (address and port), in case the flow doesn't use a transport protocol a protocol number would be used with another identifier (SPI for IPsec)

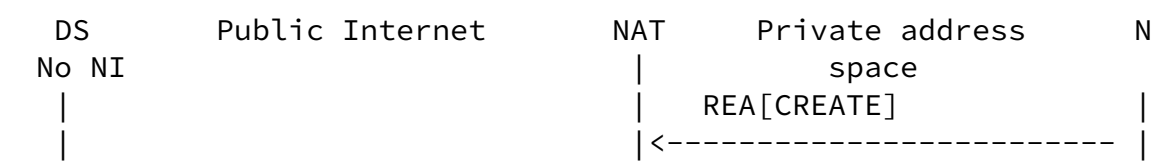
QUERY responses should provide the following information:

- List of active sessions associated to a user
- Related information to a session: session ID, flow description and policy rule state information

3.4 NATFW NSLP proxy mode of operation

3.4.1 Reserving External Addresses and triggering Create messages

Some migration scenarios need specialized support to cope with cases where only the receiving side is running NSIS. End-to-end signaling is going to fail without NSIS support at both data sender and data receiver, unless the NATFW NSLP also gives the NR the ability to install sessions. In this case, a NR can signal towards the Opportunistic Address as is done in the standard REA message handling scenario Section 3.3.2. The message is forwarded until it reaches the edge-NAT and retrieves a public IP address and port number. Unlike the standard REA message handling case no RESPONSE message is sent. Instead a CREATE message is generated by the edge-NAT. This CREATE request message is sent towards NR with DS as source address (if the source address is known, otherwise the edge NAT address is used as source address) and thereafter follows the regular processing rules as for CREATE messages sent by the NI.



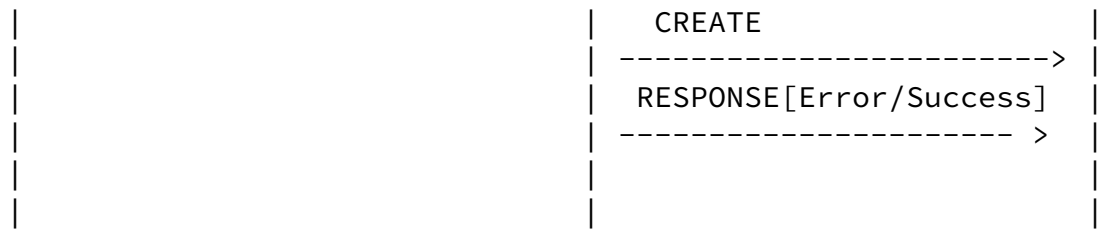


Figure 16: REA Triggering Sending of CREATE Message

This behavior requires within the REA message an indication to the edge NAT if either a RESPONSE message or a CREATE message should be

used. In addition when the CREATE message is requested (as opposed to a RESPONSE message) the REA message the data sender address. A slight variant, shown in Figure 17 , could also be handled by requesting within the REA message that a RESPONSE message needs to be sent on the existing pinned down path as well as a CREATE message on a newly discovered path between the Edge NAT and the NR. This variant would allow the handling of asymmetric routes, which could go through internal firewalls, within the local network.

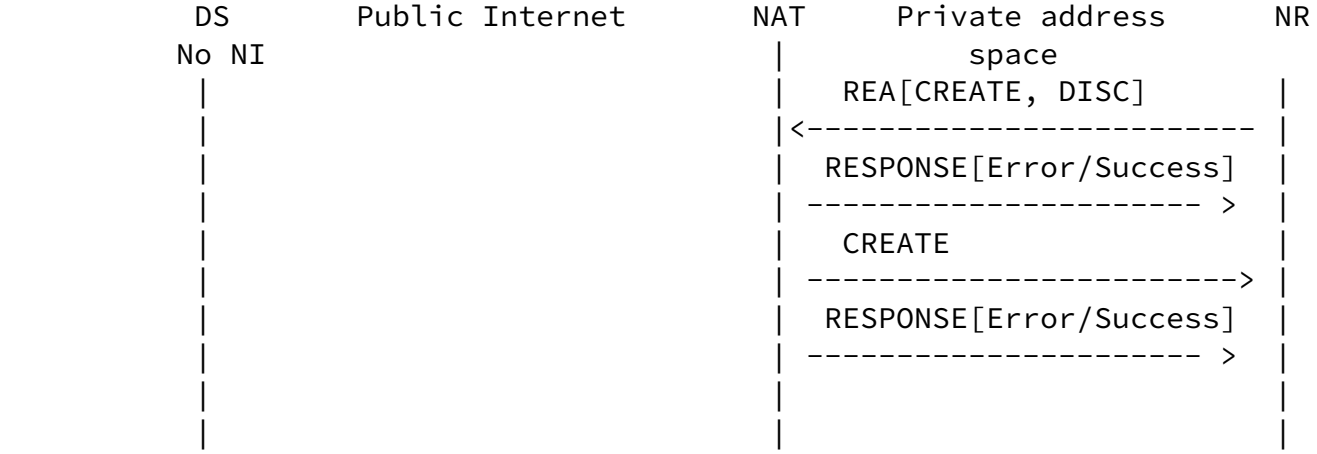
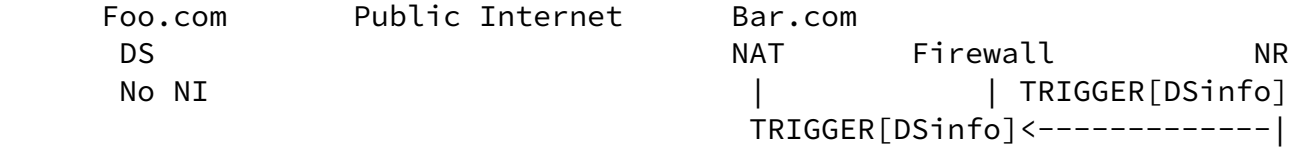


Figure 17: REA Triggering Sending of CREATE Message on Separate

In case a CREATE message is received from the far end NI and relates the installed session, that CREATE message would have precedence over the previous CREATE. The CREATE sent by the NI would allow to have a more granular policy rule as only the data sender could send data whereas in the REA triggered CREATE message any data source can send packets to the data receiver. The edge NAT is not aware of the applications context for which the CREATE messages were required. Hence it is up to the NR to inform the Edge NAT if there was a possibility to reduce the number of accepted data sources to the real data sender, as well as to inform the Edge NAT to refresh the established session.

For that purpose the NR will send TRIGGER messages, to the edge NAT that responded to the REA message. These messages are sent upon reception, from the user application, of further information on the Data Sender (either explicit information or implied information such as data sender address data reception address and same for the transport port). The TRIGGER messages would be sent periodically to

the Edge NAT that responded to the REA. The TRIGGER messages would be sent until either a CREATE message is received from the far-end or when the user application no longer needs the NSIS session. Figure 18 shows how TRIGGER messages would be used after the message sequences of Figure 16 or Figure 17. In case a CREATE message is received from the far end NI and relates to the installed session, that CREATE message would have precedence over the triggered CREATE messages. TRIGGER messages do not require to be responded back with a RESPONSE message on the existing established reverse path. The benefits of using REA triggering a CREATE and then using the TRIGGER messages are that an end-host doesnt need to know if the far-end support the NSIS protocol.



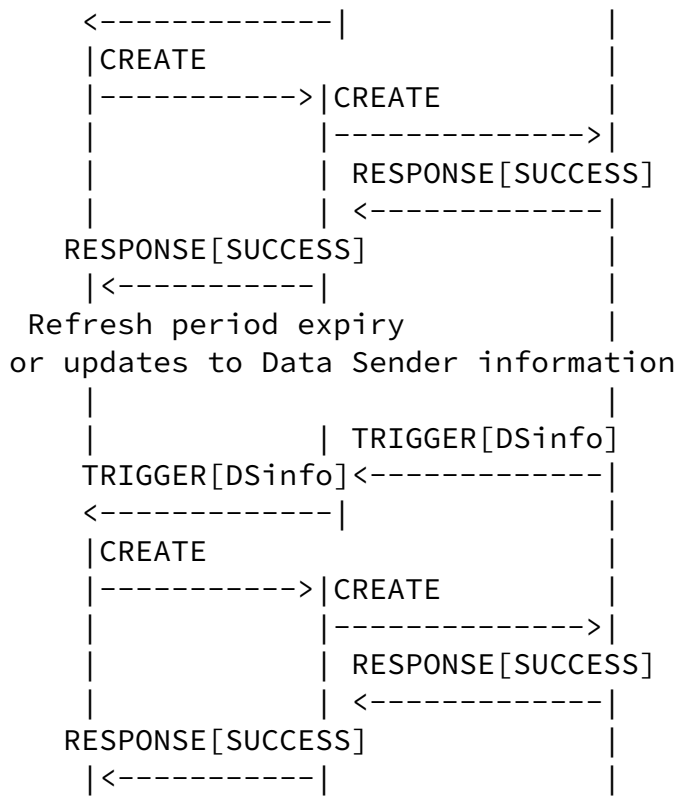


Figure 18: TRIGGER message usage

3.4.2 Using CREATE messages to Trigger Reverse Path CREATE Messages

In certain network deployments, where a NATFW NE might not be available on the end-host (Figure 19) or the NSIS messages are scoped (Figure 20) implicitly or explicitly with a scoping object, a CREATE message could be used to trigger another CREATE message sent by the last NF terminating the CREATE message. There are two options for this mode:

- The returning CREATE message could follow the established reverse path using GIMPS routing state ([3],Section 3.4.2.1)
- Trigger the GIMPS layer to discover the reverse path, which would require that the first CREATE message provides the message target address (Section 3.4.2.2).

3.4.2.1 CREATE Responses Sent on Previously Pinned Down Reverse Path

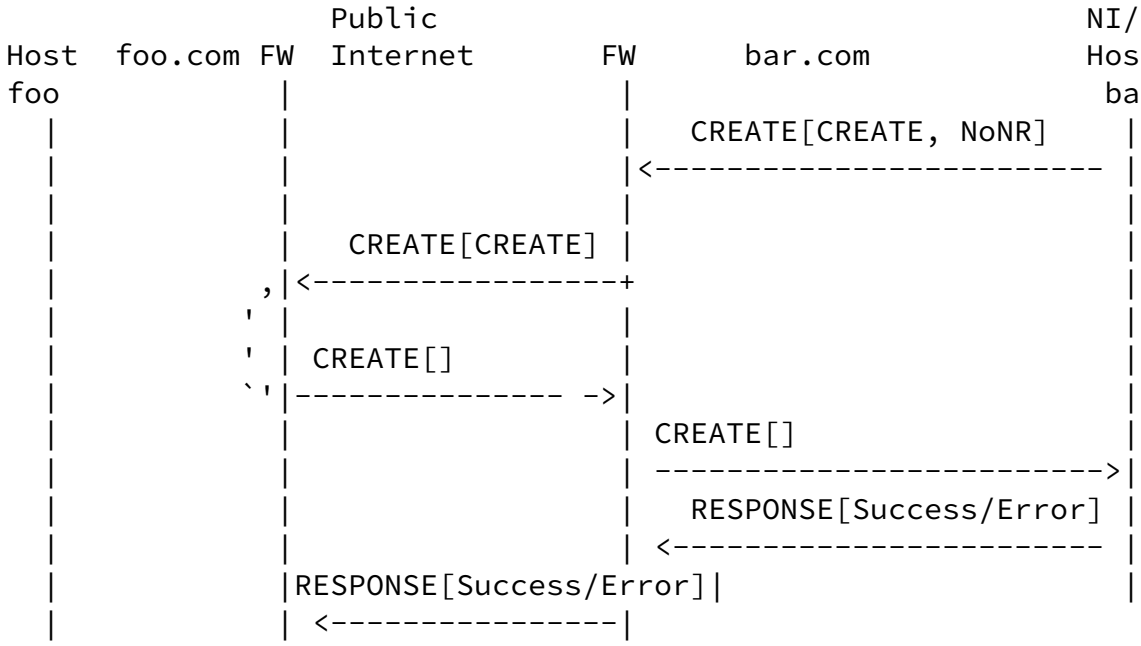


Figure 19: CREATE triggering CREATE Message Sending with no Scoping and using Existing Reverse Path State

In Figure 19, the first CREATE indicates that if the message can not reach its destination, a CREATE message should be sent back to the NI by the last reached NATFW NE. As in Section 3.4.1 this mode of operation requires that the CREATE message indicate the type of

required response which in this case is a CREATE message. However this response type is subject to a condition: only if the NR can not respond. This conditional behavior requires a specific flag to indicate it. In this example, the NI does not require that the last NATFW NF responds via a different reverse path than that already pinned down.

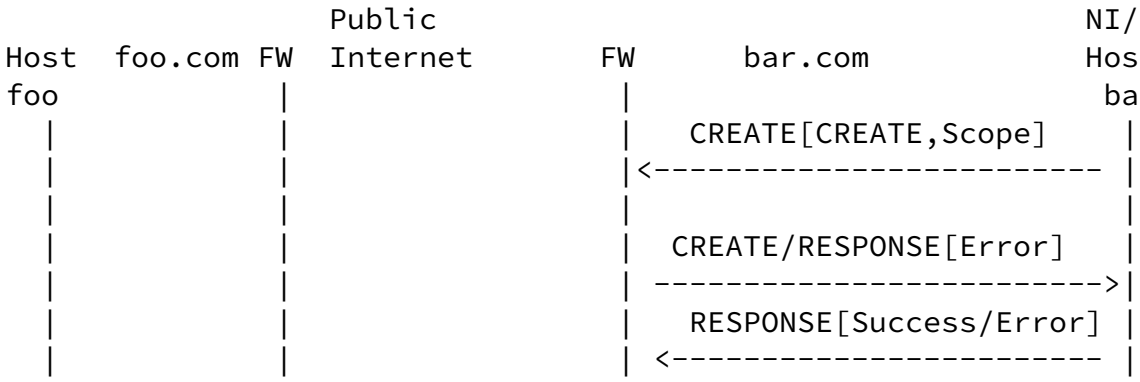


Figure 20: CREATE Triggering CREATE Message Sending with Scoping and using Existing Reverse Path State

In Figure 20, the first CREATE indicates that once the end of the scope is reached, the last NATFW NSLP will respond with a CREATE message (if the first CREATE request was successful). As in Section 3.4.1, this mode of operation requires that the CREATE message indicate the type of response required which in this case is a CREATE message. As the CREATE needs to terminate at a scope end, the scope need to be provided within the CREATE message. In this example, the NI doesnt require that the last NATFW NF responds via a different reverse path than the already pinned down.

3.4.2.2 CREATE Responses Sent on Separately Established Reverse Path

In certain network topologies, where several NATFW NSLP are deployed on alternate paths, it is better to minimize asymmetric route issues that could occur when sending the CREATE message on the existing pinned down reverse path.

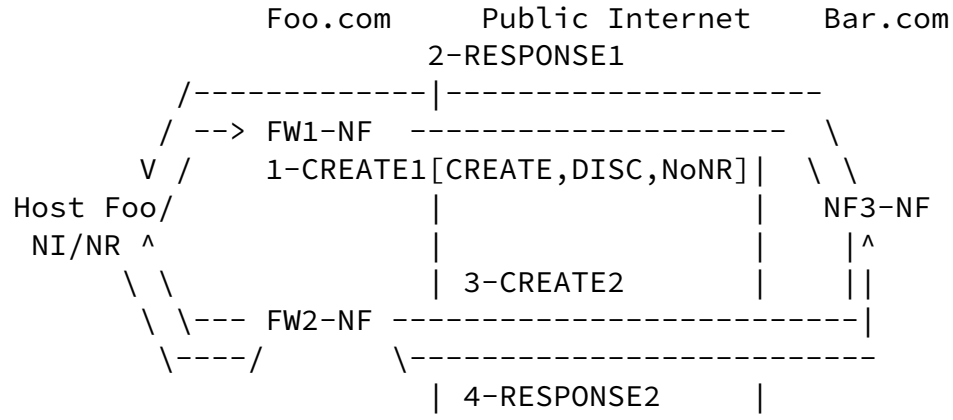


Figure 21: CREATE Triggering Sending of CREATE Message with Scoping and Using Separate Reverse path

To minimize the asymmetric route problem, the node responding with a CREATE message would request the NTLP to rediscover the reverse path. A RESPONSE message would be sent on the existing pinned down reverse path (Step 2 in Figure 21), and a CREATE would be sent on a newly discovered reverse path (Step 3 in Figure 21). Upon reception of the latter message, the initiating NI will respond with a RESPONSE message (Step 4 in Figure 21) as is done for the normal CREATE message operations (Section 3.3.1). The CREATE message would need to indicate to the last NATFW NF that a CREATE must be sent on a separately discovered path and that a RESPONSE message needs to be sent on the established pinned down reverse path. The new CREATE message need to indicate to the NI that this session is bound to the previous session. In addition the first message should indicate that the last available NATFW NF will need to terminate the message and start the above procedures (similar to Figure 19). The model could also be applied when a scope is used, instead of terminating on the last NATFW NF, the message will terminate on the end of the scope.

3.5 Calculation of Session Lifetime

NATFW NSLP sessions, and the corresponding policy rules which may have been installed, are maintained via soft-state mechanism. Each session is assigned a lifetime and the session is 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 extension refresh messages (see Section 3.3.3): this is the task of the NSIS initiator.

NSIS initiator MUST choose a session lifetime (expressed in seconds) value before sending any message (except 'delete session' messages)

to other NSLP nodes. The session lifetime value is calculated based on:

- The number of lost refresh messages to cope with
- The end to end delay between the NI and NR
- Network vulnerability due to session hijacking ([21]). Session hijacking is made easier when the NI does not remove explicitly the session.
- The user application's data exchange duration, in terms of seconds, minutes or hours and networking needs. This duration is modeled as M x R, with R the message refresh period (in seconds) and M a multiple of R.

As opposed to the NTLP Message Routing state [3] lifetime, the NSLP session lifetime doesnt require to have a small value since the NSLP state refresh is not handling routing changes but security related concerns. [14] provides a good algorithm to calculate the session lifetime as well as how to avoid refresh message synchronization within the network. [14] recommends:

1. The refresh message timer to be randomly set to a value in the range [0.5R, 1.5R].
2. To avoid premature loss of state, L (with L being the session lifetime) must satisfy L >= (K + 0.5)*1.5*R, where K is a small integer. Then in the worst case, K-1 successive messages may be lost without state being deleted. Currently K = 3 is suggested as the default. However, it may be necessary to set a larger K value for hops with high loss rate. Other algorithms could be used to define the relation between the session lifetime and the refresh message period, the provided algorithm is only listed as an example.

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 NFs processing the request message along the path MAY change the requested lifetime to fit their needs and/or local policy. If an NF changes the lifetime value it must also indicate the corresponding refresh message period. NFs MUST NOT increase the lifetime value unless the lifetime value was below their acceptable range; 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. The NSLP responder MAY also lower the requested lifetime to an acceptable value (based on its local policies). 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 within the tolerable lifetime range defined in their local policies. They MAY reject the lifetime and generate a 'lifetime not acceptable' error response message. Figure 22 shows the procedure with an example, where an initiator requests 60 seconds lifetime in the CREATE message and the lifetime is shortened along the path by the forwarder to 20 seconds and by the responder to 15 seconds.

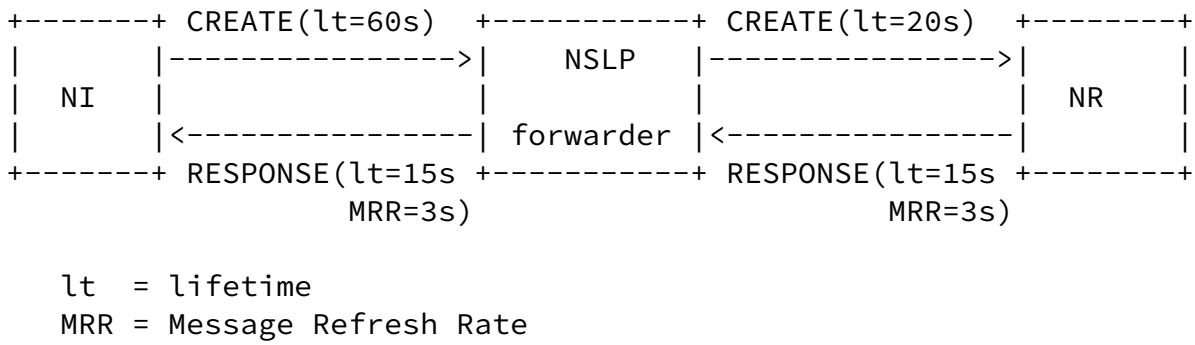


Figure 22: Lifetime Calculation Example

3.6 Middlebox Resource

TBD: 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.

3.7 De-Multiplexing at NATs

Section 3.3.2 describes how NSIS nodes behind NATs can obtain a publicly reachable IP address and port number at a NAT. The information IP address/port number can then be transmitted via a signaling protocol and/or third party to the communication partner that would like to send data towards hosts behind the NAT. 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.

The de-multiplexing method uses information stored at NATs (such as mapping of public IP address to private, transport protocol, port numbers) and information given by NTLP's flow routing information.

3.8 Selecting Opportunistic Addresses for REA

REA 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 in the first place, there is a need to choose a destination address for REA messages. This destination address can be the final target, but for applications which do not provide an upfront address, the destination address has to be chosen independently. Choosing the 'correct' destination IP address may be difficult and it is possible there is no 'right answer'. [19] shows choices for SIP and this section provides some hints about choosing a good destination IP address.

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, this could happen in a network deployment such as in Figure 12. As a consequence it might be necessary to advertise a new (and different) external IP address within the application (which may or may not allow that) 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 of 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.

4. NATFW NSLP NTLP Requirements

The NATFW NSLP requires the following capabilities from the NTLP:

- Ability to detect that the NSIS Responder does not support NATFW NSLP. This capability is key to launching the proxy mode behavior as described in Section 3.4 and [17].
- Detection of NATs and their support of the NSIS NATFW NSLP. If the NTLP discovers that the NSIS host is behind an NSIS aware NAT, the NR will send REA messages to the opportunistic address. If the NTLP discovers that the NSIS host is behind a NAT that does not support NSIS then the NSIS host will need to use a separate NAT traversal mechanism.
- Message origin authentication and message integrity protection
- Transport of information used for correlation purposes between the NSIS protocol suite and user application layers. This requirement

allows NSLP NATFW to check that the message was solicited by prior application message exchanges before an NTLP messaging association is established between an NR and the upstream NF.

- Detection of routing changes
- Protection against malicious announcement of fake path changes, this is needed to mitigate a threat discussed in section 7 of [21]

5. NATFW NSLP Message 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 aligned to 32 bit boundaries and object lengths that are not multiples of 32 bits must be padded to the next higher 32 bit

multiple.

The whole NSLP message is carried as payload of a NTLP message.

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

5.1 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 23.

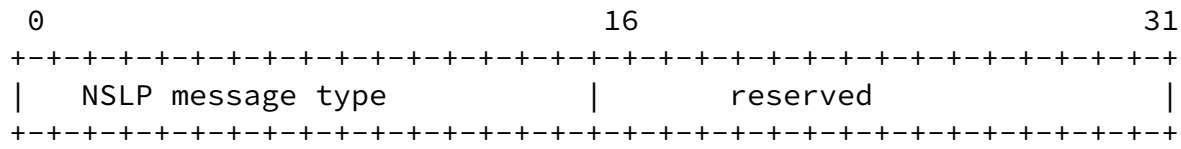


Figure 23: Common NSLP header

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

5.2 NSLP message types

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

- 0x0101 : CREATE
- 0x0102 : RESERVE-EXTERNAL-ADDRESS(REA)
- 0x0103 : QUERY
- 0x0104 : NOTIFY
- 0x0105 : RESPONSE
- 0x0106 : TRIGGER

Defined message types for responses are (TBD):

- TBD

5.3 NSLP Objects

NATFW NSLP objects use a common header format defined by Figure 24. 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.

Note that all objects MUST be padded always to 32 bits.

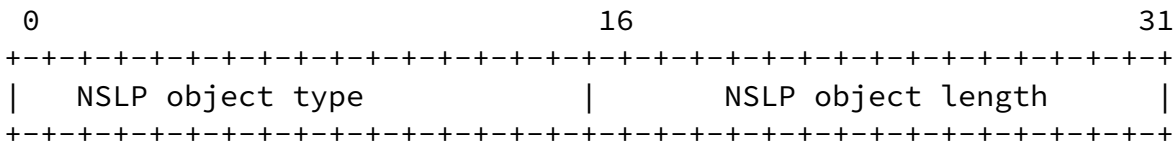


Figure 24: Common NSLP object header

The length is the total length of the object without the object header. The unit is a word, consisting of 4 bytes. The particular values of type and length for each NSLP object are listed in the subsequent sections that define the NSLP objects.

TBD: Processing of unknown options is currently subject to discussions within the working group. It is proposed to extend the NSLP object header with some bits that indicate treatment of unknown options. The compatibility bits (CP) are coded into two 2 bits and determine the action to take upon receiving an unknown option. The applied behavior based on the CP bits is:

- 00 - Abort processing and report error
- 01 - Ignore object and do not forward
- 10 - Ignore object and do forward

All other combinations MUST NOT be set and objects carrying these other CP bit combinations MUST discarded.

5.3.1 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.

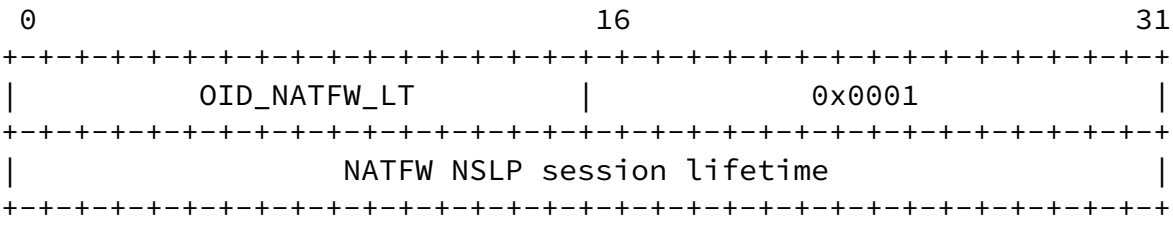


Figure 25: Lifetime object

5.3.2 External Address Object

The external address objects can be included in RESPONSE messages (Section 5.4.4) only. It contains the external IP address and port number allocated at the edge-NAT. Two fields are defined, the external IP address, and the external port number. For IPv4 the object with value OID_NATFW_IPv4 is defined. It has a length of 8 bytes and is shown in Figure 26.

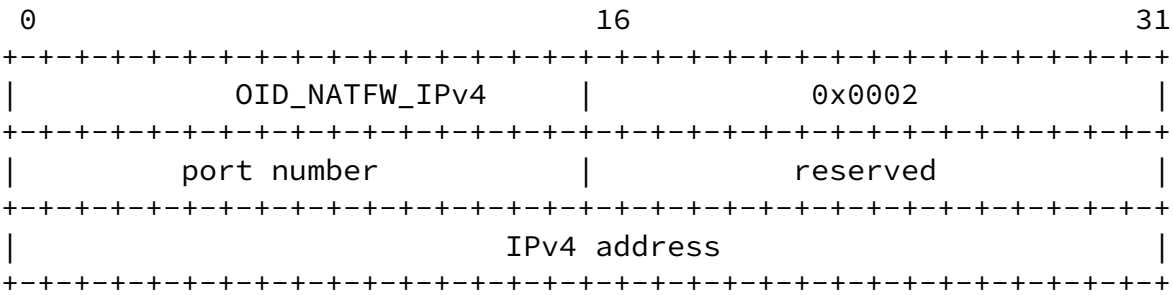


Figure 26: External Address Object for IPv4 addresses

For IPv6 the object with value OID_NATFW_IPv6 is defined. It has a length of 20 bytes and is shown in Figure 27.

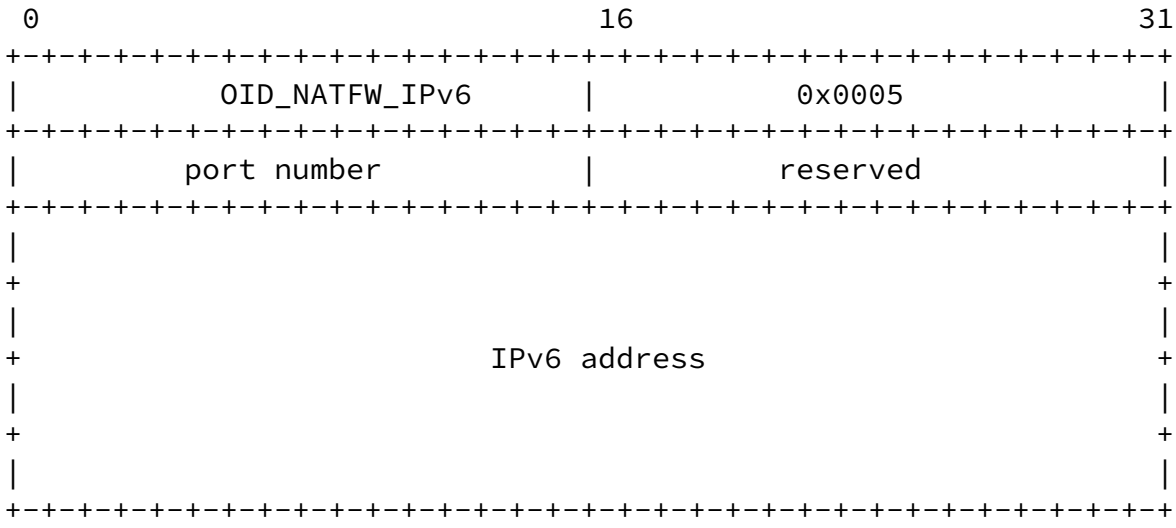


Figure 27: External Address Object for IPv6 addresses

5.3.3 Extended Flow Information Object

In general, flow information is kept at the NTLP level during signaling. The message routing information of the NTLP carries all necessary information. Nevertheless, some additional information may be required for NSLP operations. The 'extended flow information' object carries this additional information about action to be taken on the installed policy rules and subsequent numbers of policy rules.

These fields are defined for the policy rule object:

- Rule action: This field indicates the action for the policy rule to be activated. Allow values are 'allow' (0x01) and 'deny' (0x02)
- Number of ports: This field gives the number of ports that should be allocated beginning at the port given in NTLP's message routing information.

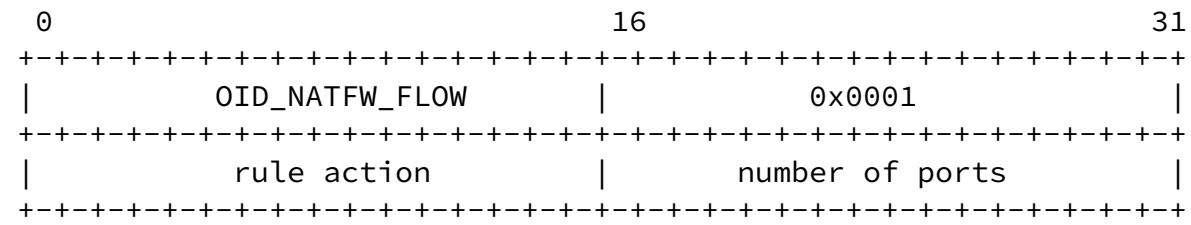


Figure 28: Extended Flow Information

5.3.4 Response Code Object

This object carries the response code, which may be indications for either a successful request or failed request depending on the value of the 'response code' field.

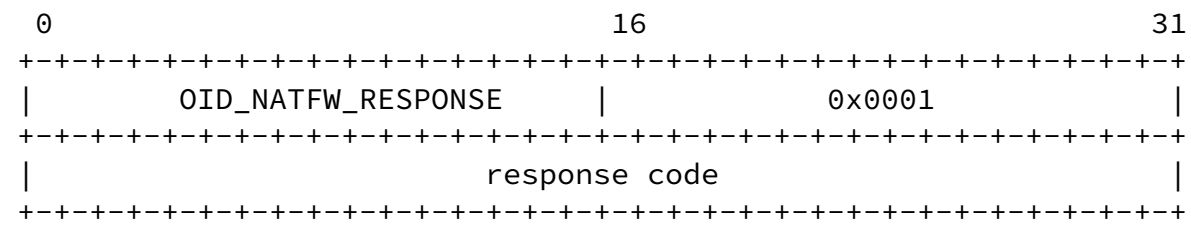


Figure 29: Response Code Object

TBD: Define response classes, success codes and error codes.

Possible error 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

5.3.5 Response Type Object

The response type object indicates that a specific response is needed to the NSLP responder.

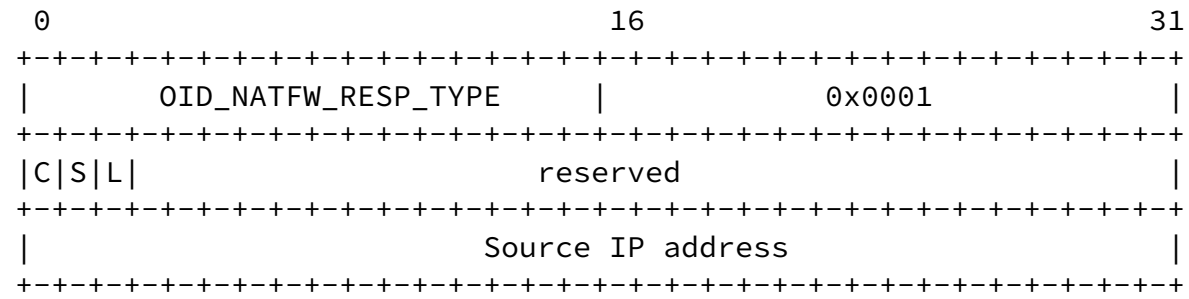


Figure 30: Response Type Object

If the C bit is set to 1 the required response is a CREATE request message, otherwise a RESPONSE message. If the S bit is set to 1 the scoping object MUST be used. If the L bit is set to 1 the CREATE request message is ONLY sent if the message does not reach its target, even though the if the C bit is set.

The source IP address is optional and may be set to a zero IP address or to a real IP address. If set to a real address, NATFW NSLP uses this address as assumed data sender's address.

5.3.6 Message Sequence Number Object

XXX Text is missing.

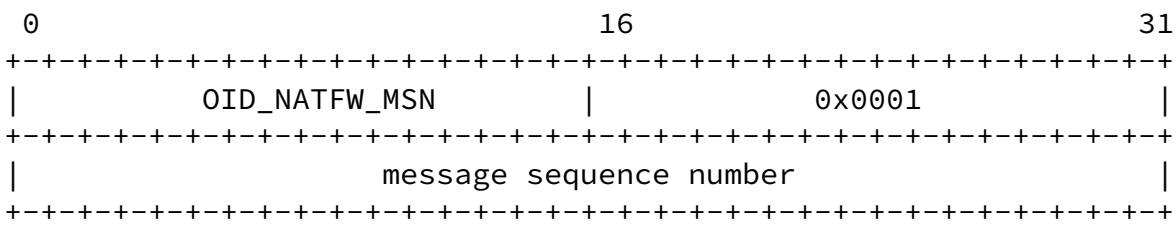


Figure 31: Message Sequence Number Object

5.3.7 Scoping Object

The scoping object determines the allowed scope for the particular message.

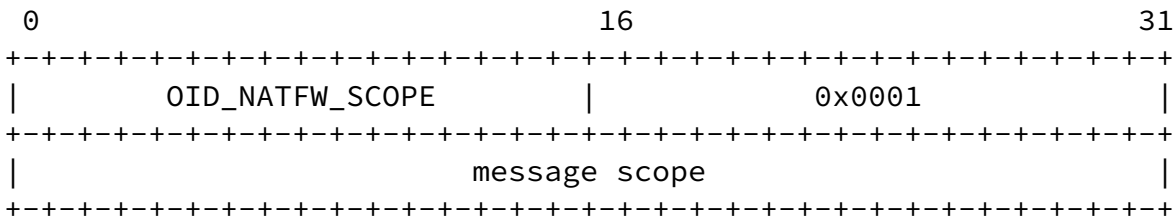


Figure 32: Scoping Object

These 'message scope' values are allowed: region, single hop.

5.3.8 Bound Session ID Object

This object carries a session ID and is used for QUERY messages only.

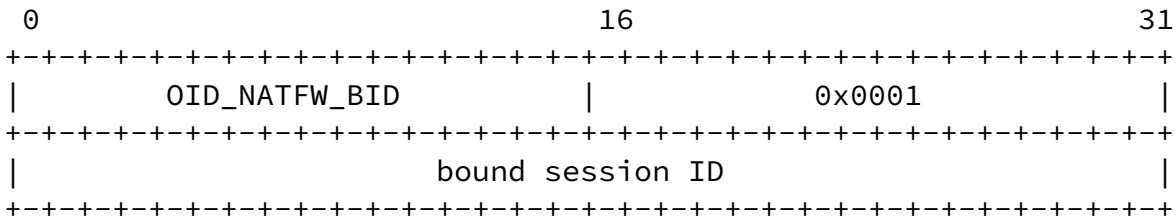


Figure 33: Bound Session ID Object

5.3.9 Notify Target Object

This object carries the IP address of the notify target node. TBD: Details on this, like IPv6 version etc.

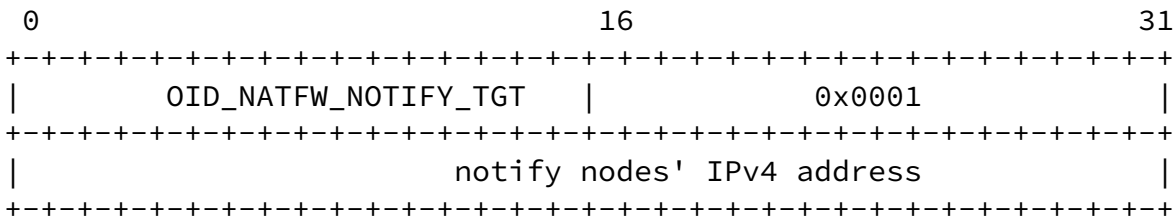


Figure 34: Notify Target Object

5.4 Message Formats

This section defines the content of each NATFW NSLP message type. The message types are defined in Section 5.2. 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. Objects are marked either with

mandatory [M] or optional [O]. Where [M] implies that this particular object MUST be included within the message and where [O] implies that this particular object is OPTIONAL within the message.

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

5.4.1 CREATE

The CREATE request message is used to create NSLP sessions and to create policy rules. Furthermore, CREATE messages are used to refresh sessions and to delete them.

The CREATE message carries these objects:
- Lifetime object [M]
- Extended flow information object [M]
- Message sequence number object [M]
- Respose type object [O]
- Scoping object[O]
- Notify target [O]

The message 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. When the CREATE messages is received by a node operating in proxy mode Section 3.4 the NI address is the NR address from the message that triggered the CREATE to be sent, if that address is not valid (wildcarded) the proxy node address is used instead. The NR address would be the NI's address provided by the message routing information of the message that triggered the CREATE.

5.4.2 RESERVE-EXTERNAL-ADDRESS (REA)

The RESERVE-EXTERNAL-ADDRESS (REA) request message is used to target 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:
- Lifetime object [M]

- Message sequence number object [M]
- Response type object [M]
- Scoping object [M]
- Extended flow information [O]

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 3.8). 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.

5.4.3 TRIGGER

The TRIGGER request message is used in proxy mode operation. XXX

The TRIGGER request message carries these objects:
- Lifetime object [M]
- Message sequence number object [M]
- Response type object [M]
- Scoping object [M]
- Extended flow information [O]

XXX

5.4.4 RESPONSE

RESPONSE messages are responses to CREATE, REA, and QUERY messages.

The RESPONSE message carries these objects:
- Lifetime object [M]
- Response object [M]
- External address object ([M] for success responses to REA)

This message is routed upstream.

5.4.5 QUERY

QUERY messages are used for diagnosis purposes.

The QUERY message carries these objects:
- Response object [M]
- Message sequence number object [M]

- Scoping object [M]
- Bound session ID [O]

This message is routed downstream.

5.4.6 NOTIFY

The NOTIFY messages is used to report asynchronous events happening along the signaled path to other NATFW NSLP nodes.

The NOTIFY message carries this object:
- Response code object with NOTIFY code [M].

The message routing information in the NTLP MUST be set to DS as source address and DR as destination address, forwarding direction is upstream (Note that Section 5.4.6 discusses some design options regarding the message transport). The session id object must be set to the corresponding session that is effected by this asynchronous event.

6. NSIS NAT and Firewall Transition 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.

7. Security Considerations

Security is of major concern particularly in case of Firewall traversal. Security threats for NSIS signaling in general have been described in [6] and they are applicable to this document. [21] extends this threat investigtion by considering NATFW NSLP specific threats. Based on the identified threats a list of security requirements have been defined. As an important requirement for security protection it is necessary to provide

- data origin authentication
- replay protection
- integrity protection and
- optionally confidentiality protection

between neighboring NATFW NSLP nodes.

To consider the aspect of authentication and key exchange we aim to reuse the mechanisms provided in [3] between neighboring nodes.

Some scenarios also demand security between non-neighboring nodes but this aspect is still in discussions. A possible commonality with the QoS NSLP has been identified and CMS [24] has been investigated as a possible candidate for security protection between non-neighboring entities. Note that this aspect also includes some more sophisticated user authentication mechanisms, as described in [23]. With regard to end-to-end security the need for a binding between an NSIS signaling session and application layer session has been described in Section 3.3 of [6].

In order to solicit feedback from the IETF community on some hard security problems for path-coupled NATFW signaling a more detailed description in [22] is available.

The NATFW NSLP is a protocol which may involve a number of NSIS nodes

and is, as such, not a two-party protocol. This fact requires more thoughts about scenarios, trust relationships and authorization mechanisms. This section lists a few scenarios relevant for security and illustrates possible trust reationships which have an impact to authorization. More problematic scenarios are described in Appendix A.

7.1 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), authorization is very important due to the nature of middleboxes.

Different types of trust relationships may affect different categories of middleboxes. As explained in [26], establishment of a financial relationship is typically very important for QoS signaling, whereas financial relationships are less directly of interest for NATFW middlebox signaling. It is therefore not particularly surprising that there are differences in the nature and level of authorization likely to be required in a QoS signaling environment and in NATFW middlebox signaling. For NATFW middlebox signaling, a stronger or weaker degree of authorization might be needed. Typically NATFW signaling requires authorization to configure and traverse particular nodes or networks which may derive indirectly from a financial relationship. This is a more 'absolute' situation either the usage is allowed or not, and the effect on both network owner and network user is 'binary'.

Different trust relationships that appear in middlebox signaling environments are described in the subsequent sub-sections. QoS signaling today uses peer-to-peer trust relationships. They are simplest kind of trust relationships. However, there are reasons to believe that this is not the only type of trust relationship found in today's networks.

7.1.1 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 (after manual configuration), or has been 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 distant nodes. 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 35 does not illustrate the trust relationship between the end host and the access network.

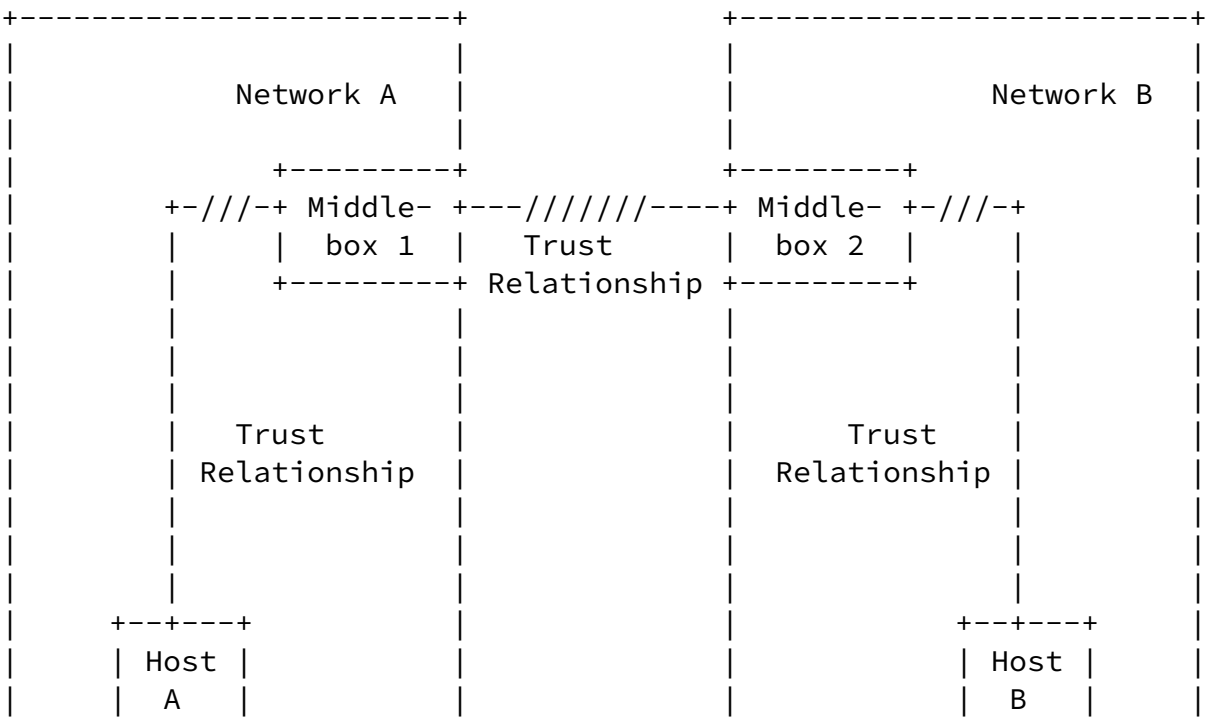
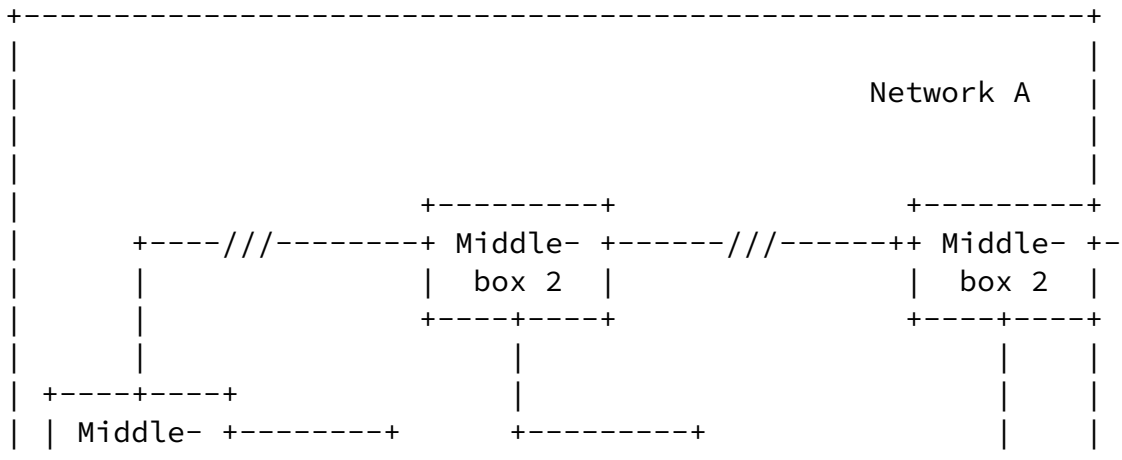


Figure 35: Peer-to-Peer Trust Relationship

7.1.2 Intra-Domain Trust Relationship

In larger corporations, often more than one middlebox is used to protect or serve 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 with 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 interactions are avoided. This provides both a performance improvement without a security disadvantage since a single administrative domain can be seen as a single entity. Figure 36 illustrates a network structure which uses a centralized entity.



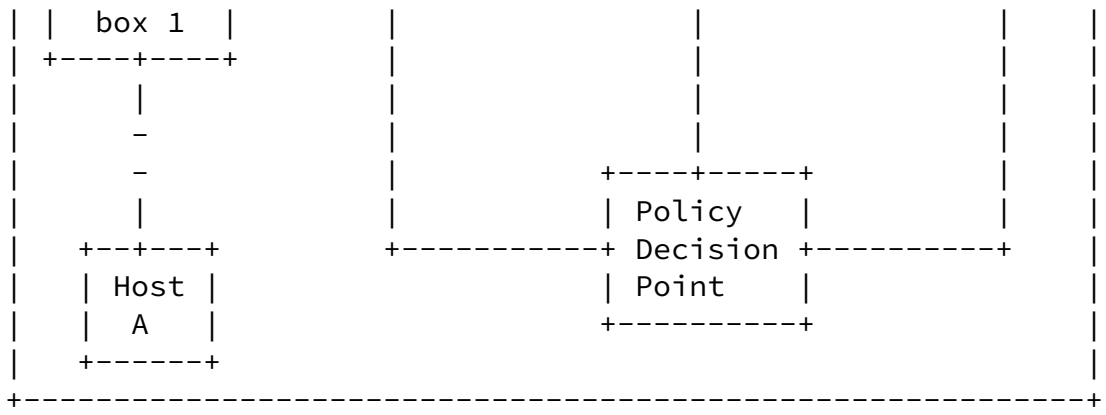
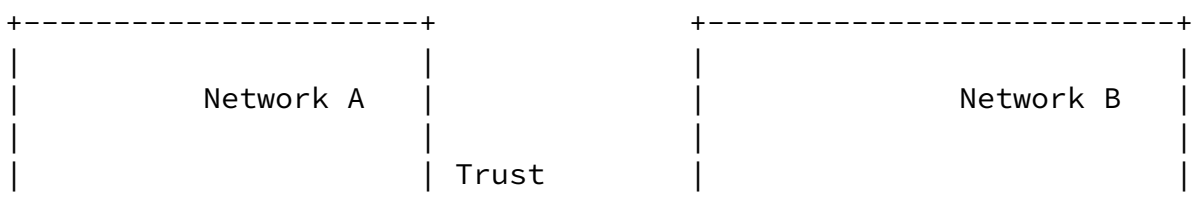


Figure 36: Intra-domain Trust Relationship

7.1.3 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 result in an error message being 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 37 shows the slightly more complex trust relationships in this scenario.



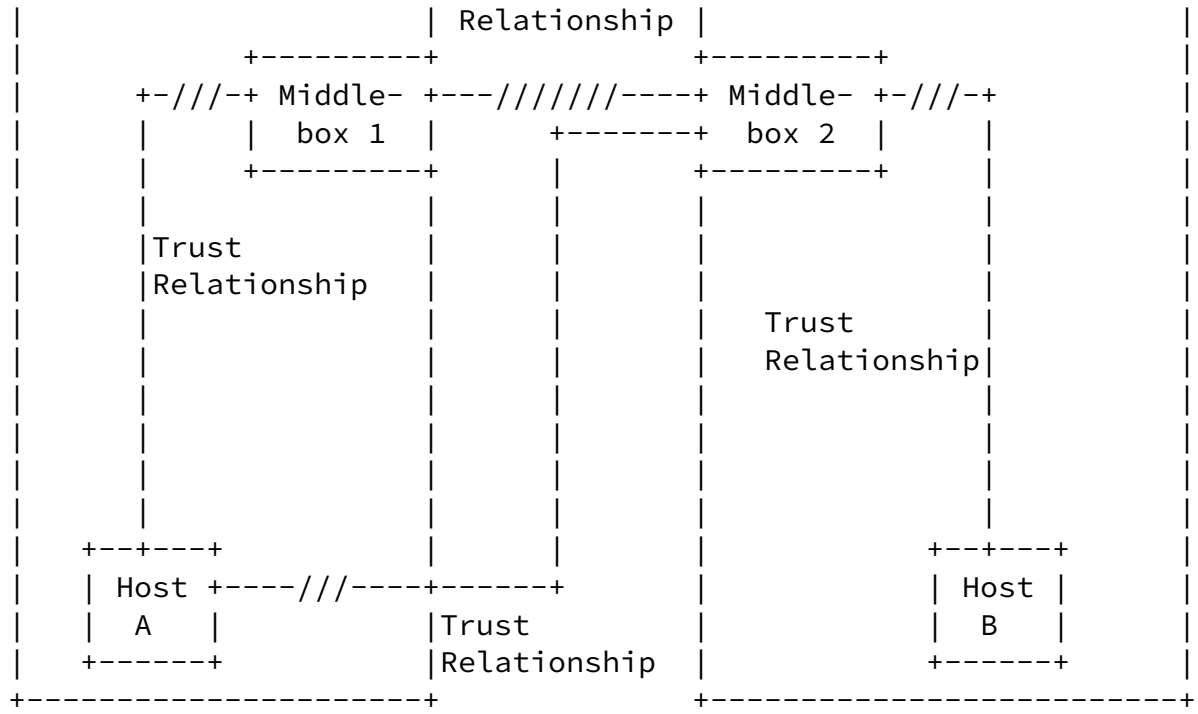


Figure 37: End-to-Middle Trust Relationship

Finally it should be noted that installing packet filters provides some security, but also 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.

8. Open Issues

The NATFW NSLP has a series of related documents discussing several other aspects of path-coupled NATFW signaling, including security [22], migration (i.e., traversal of nsis unaware NATs) [17], intra-realm signaling [18], and inter-working with SIP [19]. Summaries of the outcomes from these documents may be added, depending on WG feedback, to a later version of this draft.

A more detailed list of open issue can be found at: http://nsis.srmr.co.uk/cgi-bin/roundup.cgi/nsis-natfw-issues/index

9. 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.

10. References

10.1 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-02.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.

10.2 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, February 2004.

[21] Fessi, A., Brunner, M., Stiemerling, M., Thiruvengadam, S., Tschofenig, H. and C. Aoun, "Security Threats for the NAT/ Firewall NSLP", DRAFT draft-fessi-nsis-natfw-threats-01.txt, July 2004.

[22] Tschofenig, H., "Path-coupled NAT/Firewall Signaling Security Problems", draft-tschofenig-nsis-natfw-security-problems-00.txt (work in progress), July 2004.

[23] Tschofenig, H. and J. Kross, "Extended QoS Authorization for the QoS NSLP", draft-tschofenig-nsis-qos-ext-authz-00.txt (work in progress), July 2004.

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

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

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

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

[28] 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.

[29] 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.

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

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

[32] 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.

[33] 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.

[34] 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.

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

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

[37] Rosenberg, J., "Traversal Using Relay NAT (TURN)", draft-rosenberg-midcom-turn-04 (work in progress), February 2004.

[38] Westerinen, A., Schnizlein, J., Strassner, J., Scherling, M., Quinn, B., Herzog, S., Huynh, A., Carlson, M., Perry, J. and S. Waldbusser, "Terminology for Policy-Based Management", RFC 3198, November 2001.

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 Tschofenig
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 35, 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.

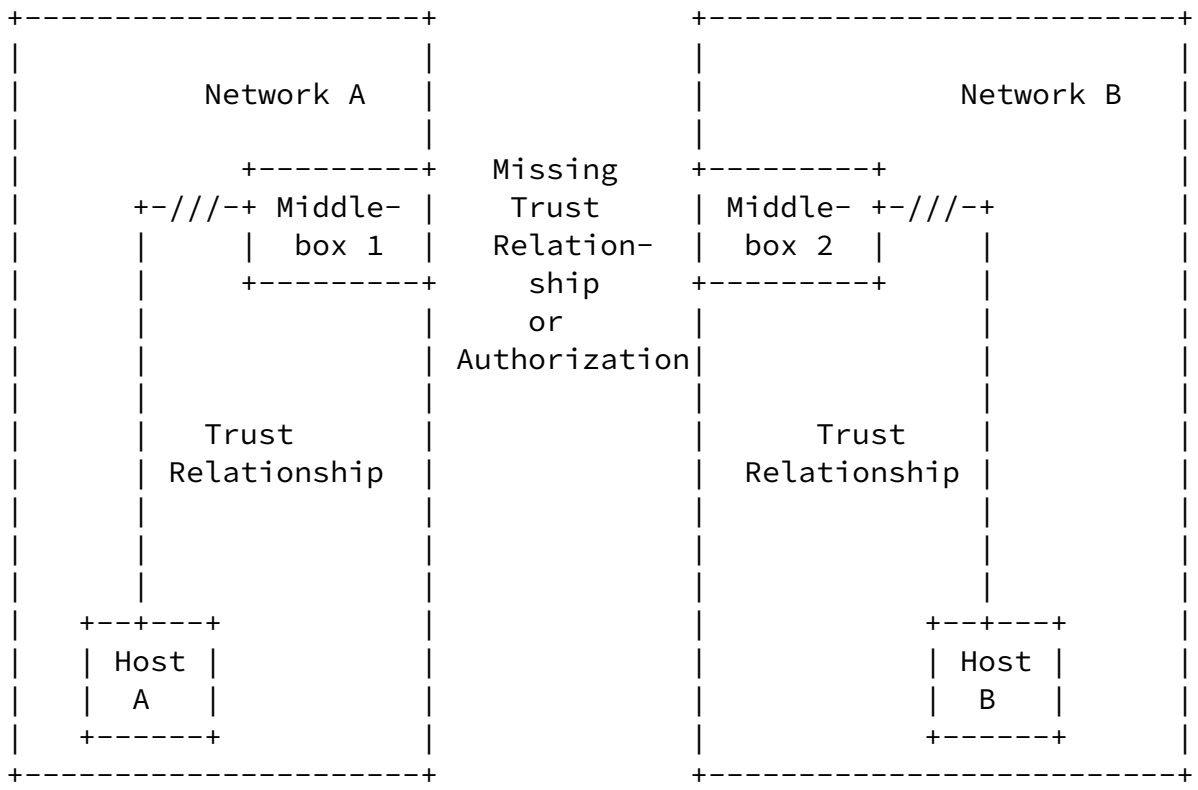


Figure 38: Missing Network-to-Network Trust Relationship

Figure 38 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 3.3.2 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 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; and Reinaldo Penno for his comments on the initial version of the document. Furthermore, we would like thank Elwyn Davis for his valuable help and input.

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights 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; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat 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 implementers or users of this

specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

Copyright Statement

Acknowledgment

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