Internet Draft

Ilya Freytsis
Cetacean Networks
Robert Hancock
Siemens/Roke Manor Research
Georgios Karagiannis
Ericsson
John Loughney
Nokia
Sven Van den Bosch
Alcatel

Document: draft-hancock-nsis-fw-00.txt
Expires: December 2002 June 2002

Next Steps in Signaling: A Framework Proposal

Status of this Memo

Abstract

The NSIS working group is considering protocol developments in signaling for resources for a traffic flow along its path in the network. The requirements for such signaling are being developed in a separate document [2]; This Internet Draft proposes a framework for such signaling. This initial version provides a model for describing the entities that take part in the signaling and the ways in which they can be used in different modes of operation. It also discusses the overall structure of such a signaling protocol. Finally, it considers the possible interactions of NSIS signaling with other protocols and functions, including security issues.

Conventions used in this document

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 [3].

Table of Contents

1. Introduction

NSIS will work on signaling from an end point that follows a path through the net that is determined by layer 3 routing and is used to convey information to the devices the signals pass through - the signaling can, for example, install soft state in the devices it passes through. A signaling end point could be a device along the path, which signals for a data flow that passes through it.

The intention is to allow for the NSIS protocol to be deployed in different parts of the Internet, for different needs, without requiring a complete end-to-end deployment.

There is no requirement that the per-flow information be QoS related. NSIS should only worry about how to do the signaling - what the signaling conveys should be opaque to NSIS. This document discusses 'where' the signaling takes place, with some discussion on 'how' the signaling can be done.

1.1 Scope of this Document

The scope of this document is to provide a framework for where a NSIS protocol can be used and deployed. It is not intended that NSIS will provide an over-arching architecture for carrying out resource management in the Internet. It is not intended to be used as a protocol design document.

The framework is not about what NSIS should do but how it should do it. It is not intended that this places requirements on a future NSIS

protocol. The document discusses important protocol considerations, such as mobility, security, interworking with resource management (in a broad sense). Discussions about existing signaling and resource protocols are assumed to be contained in a separate analysis document.

The initial draft of this document is more about discussing the important issues and gaining some scoping on the problem space. Future revisions will have more concrete proposals.

The purpose of this document is to develop the realms, domains and modes of operation where an NSIS protocol can be used; identify the relationship of an NSIS protocol to other protocols; and identify areas for future work.

2. Terminology

Classifier - an entity which selects packets based on the content of packet headers according to defined rules.

Interdomain traffic - Traffic that passes from one NSIS domain to another.

NSIS Domain (ND) - Administrative domain where an NSIS protocol signals for a resource or set of resources.

NSIS Entity (NE) - the function within a node which implements an NSIS protocol.

NSIS Forwarder (NF) - NSIS Entity on the path between a NI and NR which may interact with local resource management function (RMF) for this purpose. NSIS Forwarder also propagates NSIS signaling further through the network.

NSIS Initiator (NI) - NSIS Entity that initiates NSIS signaling for a network resource.

NSIS Responder (NR) - NSIS Entity that terminates NSIS signaling and can optionally interact with applications as well.

Peer session - signaling relationship between two adjacent NSIS entities (i.e. NEs with no other NEs between them).

Resource - something of value in a network infrastructure to which rules or policy criteria are first applied before access is granted. Examples of resources include the buffers in a router and bandwidth on an interface.

Resource Management Function (RMF) - an abstract concept, representing the management of resources in a domain or a node.

Service Level Agreement (SLA) - a service contract between a customer and a service provider that specifies the forwarding service a customer should receive.

Traffic characteristic - a description of the temporal behavior or a description of the attributes of a given traffic flow or traffic aggregate.

Traffic flow - a stream of packets between two end-points that can be characterized in a certain way.

3. Overall Framework Structure

3.1 Basic Signaling Entities and Interfaces

3.1.1 NSIS Entities

The NSIS protocol is intended to be used as a signaling control plane for the variety of network resources required for data traffic across the Internet. The most common examples are QoS resources, firewalls and NATs resources, etc. The NSIS signaling itself does not depend on the type of the network resources it is used for but the information it carries does. This section discusses the basic signaling entities of the protocol as well as interfaces between them.

We can identify three different roles in the NSIS signaling for resources: initiator, forwarder and responder.

The NSIS Initiator (NI) is an entity that initiates NSIS signaling (request) for the network resource. The NSIS initiator can be triggered by the different "sources" - applications, an instance of NSIS Forwarder, other protocols, network management etc. - that need network resources for a data flow. For the purpose of the NSIS discussion all these sources can be called applications. The NSIS initiator can provide feedback information to the triggering application in respect to the requested network resources. The NSIS initiator uses NSIS signaling to interact with other NSIS entities (NFs and NRs).

The NSIS Forwarder (NF) is an entity that services NSIS resource requests from NSIS initiators and other NSIS forwarders. It may interact with local resource management function (RMF). How and if this interaction takes place depends on the deployed resource management mechanism and the specific role of the NF. The NSIS forwarder propagates NSIS signaling further through the network.

The NSIS Responder (NR) is an entity that terminates NSIS signaling and can optionally interact with applications as well e.g. for the purpose of notification when network resources get allocated etc.

The signaling relationship between two NSIS entities (with no other NSIS entities between them) is called a 'Peer-session'. This concept might loosely be described as an 'NSIS hop'; however, there is no implication that it corresponds to a single IP hop.

Figure 1 depicts simplified interactions/interfaces between NI, NFs and NR as well as applications and RMFs. Note that the NI and NR could also interact with an RMF; additionally, this could be modeled as co-location of NI&NF and NR&NF. This distinction should have no impact on the operation of the protocol. Also, there is no bar on placing an NI or NR in the interior of the network, to initiate and terminate NSIS signaling independently of the ultimate endpoints of the end to end flow, and NI and NR do not have to talk via intervening NFs. An example of NSIS being used in this way is given in section 5.5.

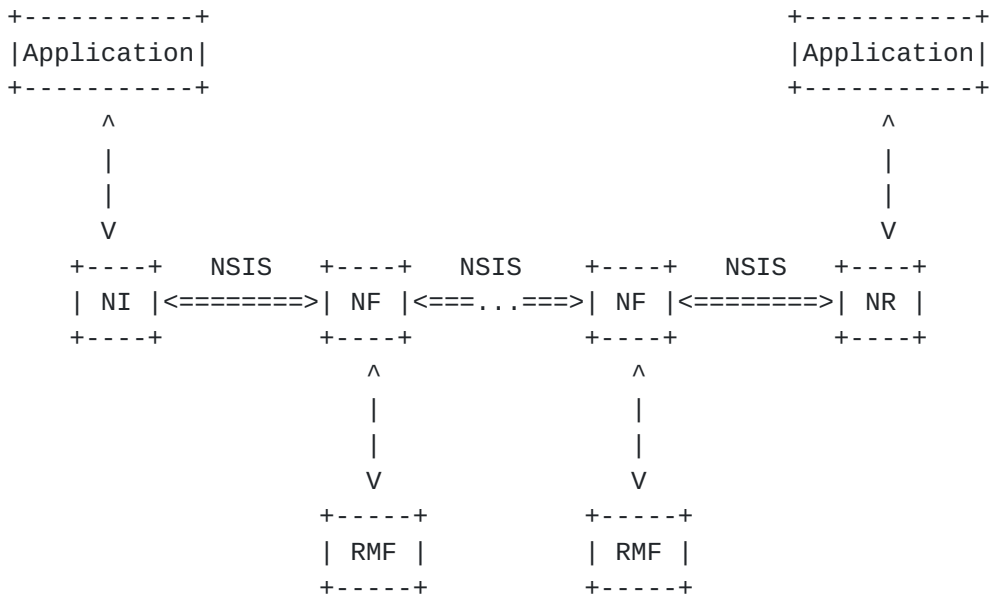

Figure 1: Basic NI/NF/NR Relationships

3.1.2 Placement of NSIS entities

The NI, NF and NR definitions do not make any assumptions about placements of NSIS signaling entities in respect to the particular part of the network or data-forwarding path.

They can be located along the data path (hosts generating and receiving data flows, edge routers, intermediate routers etc.) but it may not be the only one desirable location.

In some cases it is desired to be able to initiate and/or terminate NSIS signaling not from the end host that generates/receives the data flow but from the some other entities on the network that can be called application or service NSIS proxies. There could be various reasons for this: signaling in behalf of the end hosts that are not enabled with NSIS, consolidation of the customer accounting (authentication, authorization) in respect to consumed application and transport resources, security considerations, limitation of the physical connection between host and network etc. The proxy can communicate the relevant information to the host in the application/service specific, maybe compressed, form.

Support for NSIS proxies affects the protocol in the following way:
* The protocol should accommodate signaling with the scope of a single NSIS peer-session; the signaling could be propagated over multiple peer-sessions all the way toward the destination (end-to-end).
* In the particular case where the proxy is not on the data path, NSIS might have to be extended to allow separated data and signaling paths, although this analysis is not initially in scope.

The further discussion of these issues is given in sections 3.2.1 and 3.3.3.

As it can be seen from the usage cases presented in the NSIS requirements draft [2] the NSIS signaling procedures may depend on the part/type of the network where NSIS is used. In fact to satisfy sometimes-conflicting requirements in [2], different procedures and possibly different kinds of the NSIS protocol can be used on different parts/types of the network. Sections 3.2 and 5.5 provide more details on this topic.

3.2 Modes of Operation

This section discusses several modes of NSIS protocol operation. Each mode of NSIS operation is briefly introduced and where needed analyzed and compared with other modes of NSIS operation.

3.2.1 In-Band and Out-of-Band Signaling

In-band signaling means that the path followed by the user data packets is the same as the path followed by signaling messages. In other words, the signaling and data paths are identical. Out-of-band signaling means that the path followed by signaling messages might be different from the path used by the user data packets.

There are potentially significant differences in the way that the in and out of band signaling paradigms should be analyzed, for example in terms of scaling behavior, failure recovery, security properties, mechanism for NSIS peer discovery, and so on. These differences might or might not cause changes in the way that the NSIS protocol operates. The initial goal of NSIS and this framework is to concentrate mainly on the in-band case.

3.2.2 Inter-domain and Intra-domain Signaling

Inter-domain NSIS signaling is where the NSIS signaling messages are originated in one NSIS domain and are terminated in another NSIS domain.

In the case of in-band signaling, inter-domain NSIS signaling can be used to signal NSIS information to the edge nodes of one or more NSIS domains.

In the case of out-of-band signaling, inter-domain NSIS signaling can be used to signal NSIS information to entities that are not on the data path (i.e., "out-of-band" NFs), and additionally to signal from off-path entities to on-path edge nodes .

NSIS inter-domain signaling has to fulfill several requirements, such as:
* Basic functionality, such as scalable, simple and fast signaling. Because different networks have different resource management characteristics, such as cost of bandwidth and performance, this basic functionality may differ from one NSIS domain to another.
* All other requirements specified in [2].

Intra-domain NSIS signaling is where the NSIS signaling messages are originated, processed and terminated within the same NSIS domain. Note that these messages could be handled within a local instance of NSIS signaling; another possibility could be to piggyback them on inter-domain NSIS messages.

Intra-domain signaling can be used to signal NSIS information to the edge nodes (i.e., routers located at the border of the NSIS domain)

and to the interior nodes (i.e., routers located within the NSIS domain that are not edge nodes).

The NSIS intra-domain signaling approach has to fulfill fewer requirements than inter-domain signaling. These are:
* Basic functionality, such as scalable, simple and fast signaling. Due to the fact that different networks have different resource management characteristics, this basic functionality may differ from one NSIS domain to another.
* Provides the necessary functionality to interact between inter-domain signaling and intra-domain signaling.

3.2.3 End-to-End, Edge-to-Edge, and End-to-Edge

End-to-end: When used end-to-end, the NSIS protocol is initiated by an end host and is terminated by another end host. In this context, NSIS can be applied as needed within all of the NSIS domains between the end hosts. In the end-to-end path, NSIS may be used both for intra-domain NSIS signaling, as well as for inter-domain signaling.

Edge-to-edge: In this scenario the NSIS protocol is initiated by an edge node of a NSIS domain and is terminated by another edge node of the same (or possibly different) NSIS domain. NSIS can be applied either within one single NSIS domain, which is denoted as edge-to-edge in a single domain, or within a concatenated number of NSIS domains, which is denoted as edge-to-edge in a multi-domain. When an appropriate security trust relation exists between two or more concatenated NSIS domains, these concatenated NSIS domains are considered, in terms of NSIS, to be a single, larger NSIS domain.

End-to-edge: In this scenario the NSIS protocol is either initiated by an end host and is terminated by an edge node or is initiated by an edge node and is terminated by an end host. When using in-band signaling, the edge node may be a proxy that is located on a boundary node of a NSIS domain. If using out-of-band signaling, the edge node may be a proxy that is located on an out-of-band node that controls, or is associated with, a NSIS domain.

3.2.4 Global and Local Operation

It is likely that the appropriate way to describe the resources NSIS is signaling for will vary from one part of the network to another. In particular, resource descriptions that are valid for inter-domain links will probably be different from those useful for intra-domain operation (and the latter will differ from one NSIS domain to another).

One way to describe this issue is to consider the resource description objects carried by NSIS as divided in globally-understood objects ("global objects") and locally-understood objects ("local objects"). The local objects are only applicable for intra-domain signaling, while the global objects are mainly used in inter-domain signaling.

The purpose of this division is to provide additional flexibility in defining the objects carried by the NSIS protocol such that only those objects that are applicable in a particular setting are used. An example approach for reflecting the distinction in the signaling is that local objects could be put into separate local messages that are initiated and terminated within one single QoS (NSIS) domain and/or they could be "stacked" within the NSIS messages that are used for inter-domain signaling. These possibilities will be considered further during the protocol design activity.

3.2.5 Multicast versus Unicast

Multicast support, compared to unicast support, would introduce a level of complexity into the NSIS protocol mainly related to:
* complex state maintenance to support dynamic membership changes in the multicast groups, such as reservation state merging and maintenance.
* a state per flow has to be maintained that is used during backward routing.

3.2.6 Sender versus Receiver Initiated Signaling

A sender-initiated approach is when the sender of the data flow initiates and maintains the resource reservation used for that flow. In a receiver-initiated approach the receiver of the data flow initiates and maintains the resource reservation used for the data flow.

In the case of in-band signaling, in the sender initiated case, the sender of the data is the NSIS Initiator, while the receiver of the data is the NSIS Responder. In the receiver initiated case, receiver of the data is the NSIS Initiator, while the sender of the data is the NSIS Responder. In the case of out-band signaling, the mapping is not necessarily clear cut (for example, if the NI and NR are not located at the end systems themselves).

The main differences between the sender-initiated and receiver-initiated approaches are the following:
* Compared with the receiver-initiated approach, a sender using a sender-initiated approach can be informed faster when the reservation request is rejected. In other words, when using a sender-initiated

approach, the reservation request response time can be shorter in the case of an unsuccessful reservation than with a receiver-initiated approach.
* In a receiver-initiated approach, the signaling messages traveling from the receiver to the sender must be backward routed such that they follow exactly the same path as was followed by the signaling messages belonging to the same flow traveling from the sender to the receiver. This implies that a backward routing state per flow must be maintained. When using a sender-initiated approach, provided acknowledgements and notifications can be securely delivered to the sending node, backward routing is not necessary, and nodes do not have to maintain backward routing states.
* In a sender-initiated approach, a mobile node can initiate a reservation as soon as it has moved to another roaming subnetwork. In a receiver-initiated approach, a mobile node has to inform the receiver about its handover procedure, thus allowing the receiver to initiate a reservation.

3.2.7 Uni-Directional and Bi-Directional Reservations

It is possible that a resource will only be required for one direction of traffic, for example for a media stream with no feedback channel. Reservations for both directions of traffic may be required for other applications, for example a voice call. Therefore, the NSIS signaling protocol must allow for these uni-directional resource reservations and for bi-directional resource reservations is required.

The most basic method for bi-directional reservations is based on combining two uni-directional reservations. This means that the signaling messages from the sender of the bi-directional reservation towards a receiver are able to follow a different path from messages traveling in the opposite direction, which is necessary for on-path signaling in the presence of asymmetric routing. (Other more integrated approaches may be possible in constrained network topologies.) The bi-directional reservations can, for example, be used to make the NSIS signaling procedure required after a handover procedure more efficient.

3.3 Basic Assumptions and Critical Issues

3.3.1 Overview of Open Items and Critical Issues

Some of these issues are specific to another section of this document; for clarity and to provide an overview, these are summarized here. The subsequent subsections describe more generic assumptions and issues.

- the solution developed by NSIS must be sufficiently flexible and modular that it can be efficiently deployed and used with functionality appropriate to the part/type of the network. (Sections 3.2.2 and 3.2.3.)

- the protocol developed by the NSIS working group will operate in-band (the signaling and data paths are identical). Considerations related to a potential out-of-band solution are part of this framework, because they are also needed in order to co-exist with existing solutions. The NSIS working group currently has no plans to develop an out-of-band signaling protocol. (Section 3.2.1.)

- multicast support introduces a level of complexity into the NSIS protocol that is not needed in support of unicast applications. Therefore, a working assumption is be that the NSIS protocol should be optimized for unicast. (Section 3.2.5.)

- the NSIS protocol can be used for setup of both uni-directional and bi-directional reservations. (Section 3.2.7.)

- to function as part of a complete system, the NSIS protocol may need to be supported by extensions to other protocols. These extensions are still to be identified. (Section 4.2.)

- the NSIS protocol could be constructed on the services offered by lower layer protocols, but the dividing line between NSIS and these lower layers is not fixed. Use of standard lower layer protocols may be difficult if 'end-to-end addressing' (see section 3.3.4) is used. (Section 4.3.1.)

- it is commonly expected that a future resource signaling protocol would need to use abstract reservation identifiers. However, the precise properties needed of these identifiers are unclear, and enabling their secure use may be hard. (Sections 4.5.2 and 5.3.2.)

- use of some routing techniques (e.g. load sharing or QoS routing), even in remote parts of the network, could be incompatible with naive use of end-to-end addressing. (Sections 5.2.1 and 5.2.2.)

- the correct flow identification semantics need to be defined in the case where mobility encapsulations might make it ambiguous which addresses to use. (Section 5.3.1.)

- the interactions between mobility and resource signaling during path updating need to be further analyzed, especially from the point of view of combined overall latency. (Section 5.3.2 and 5.3.3.)

3.3.2 NI, NF, NR functionality

The basic functions that can be fulfilled by an NSIS entity are request, accept, notify, modify and release of a reservation. At this point, it is not clear which responsibilities can be assumed by each of the NSIS entities. More in particular, it is not clear whether:
- an NF can request, modify or release a reservation. If it cannot, it needs to notify the NI in order to perform these functions.
- an NR can modify and release a reservation. Even if the NR can reject or accept the reservation with modification, it might still be required to notify the NI to signal the release or modification.

3.3.3 NI, NF, NR relationship

An important open issue is related to the way in which NSIS entities maintain relations between each other. These relations could be purely local, where an NSIS entity only maintains relations with its direct neighbors. In that case, messages will be sent to and accepted from these neighbors only. Alternatively, the relations between NSIS entities could have a more global scope.

The type of NSIS peering relations may have an impact on the complexity involved with protocol security. In case of inter-domain signaling, the security relations are likely to be built between neighboring NSIS entities only for scalability reasons. In that case, each NSIS entity will establish and maintain a security relation with each of its peers and accept only messages from these peers. Conversely, there may exist larger domains of NSIS entities that have a trust relationship (trusted domains). This may be the case for intra-domain signaling. In this case, an NE may accept messages from all other NSIS entities in the domain. Both alternatives need not be mutually exclusive. It is conceivable that different instances of the NSIS protocol (or different NSIS protocols) use the NSIS security model to a larger or lesser extent, provided that overall security is not impacted. A detailed analysis of NSIS threats is available from [4].

The NSIS peering relations may also have an impact on the required amount of state at each NSIS entity. When direct interaction with remote NSIS peers is not allowed, it may be required to keep track of the path that an NSIS message has followed through the network. This can be achieved by keeping per-flow state at the NSIS entities or by maintaining a record route object in the NSIS messages.

3.3.4 NSIS Addressing

The are potentially two ways to establish a signaling connection by means of the NSIS protocol. On the one hand, the NSIS message could

be addressed to a neighboring NSIS entity (NE) that is known to be closer to the destination NE. On the other hand, the NSIS message could be addressed to the destination NE directly. We denote the latter approach as end-to-end addressing and the former as peer-session addressing.

With peer-session addressing, an NE will determine the address of the next NE based on the payload of the NSIS message (and potentially also on the previous NE). This requires the address of the destination NE to be derivable from information present in the payload. This can be achieved through the availability of a local routing table or through participation in the routing protocol. Peer-session addressing inherently supports tunneling of NSIS signaling messages between NEs, and is equally applicable to on or off path signaling.

In case of end-to-end addressing, the NSIS message will be sent with the address of the NR, which then necessarily needs to be on the data path. This requires (some of) the data-path entities to be upgraded (NSIS-aware) in order to be able to intercept the NSIS messages. The routing of the NSIS signaling follows exactly the same path as the data flow for which the reservation is requested.

3.3.5 Service description

Although the service specific part of the NSIS message is outside of the scope of the NSIS working group, it may be necessary to make some assumptions about its content in order to determine whether similar functionality needs to be foreseen in the NSIS-specific part of the message:
- It is assumed that the service description will handle pre-emption and survivability issues. These are seen as a part of the offered service and need not be present in the NSIS control layer.
- It is assumed that some flow description information is part of the NSIS control layer (see section 4.3.1 and 4.5.1). This might be needed by service-unaware entities located at address boundaries. It is not clear to which level of complexity, the flow description needs to be available at this level.
- It is not assumed that the content of the service description is independent of the NSIS control layer. It seems appropriate to allow the content of the service description to be dependent on the type of message that is sent (request/response/refresh).

3.3.6 NSIS Acknowledgement and Notification Semantics

The semantics of the acknowledgement and notification messages are of particular importance. An NE sending a message can assume responsibility for the entire downstream chain of NEs, indicating for

instance the availability of reserved resources for the entire downstream path. Alternatively, the message could have a more local meaning, indicating for instance that a certain failure or degradation occurred at a particular NSIS entity.

4. Protocol Components

4.1 Lower Layer Interfaces

Within a signaling entity, NSIS interacts with the 'lower layers' of the protocol stack for two nearly independent purposes: sending and receiving signaling messages; and configuring the operation of the lower layers themselves.

For sending and receiving messages, this framework places the lower boundary of the NSIS protocol at the IP layer. (It is possible that NSIS could use a standard transport protocol above the IP layer to provide some of its functionality; this is discussed in section 4.3.1.) The interface with the lower layers is therefore very simple:
*) NSIS sends raw IP packets
*) NSIS receives raw IP packets. In the case of peer-session addressing, they have been addressed directly to it. In the case of end-to-end addressing, this will be by intercepting packets that have been marked in some special way (by special protocol number or by some option interpreted within the IP layer, such as the Router Alert option [5] and [6].)

NSIS needs to have some information about the link and IP layer configuration of the local networking stack. For example, NSIS needs to know about:
*) [in general] how to select the outgoing interface for a signaling message, in case this needs to match the interface that will be used by the corresponding flow. This might be as simple as just allowing the IP layer to handle the message using its own routing table.
*) [in the case of IPv6] what address scopes are associated with the interfaces that messages are sent and received on (to interpret scoped addresses in flow identification, if these are to be allowed).

The way in which NSIS actually configures the lower layers to handle the flow depends on the particular NSIS application; for example, if NSIS is being used for QoS signaling, this might involve configuration of traffic classification and conditioning parameters, for example local packet queues, type of filters, type of scheduling, and so on. However, none of this is directly related to the NSIS protocol itself; therefore, this interaction is handled indirectly via a resource management function, as described in section 5.1.

4.2 Upper Layer Services

NSIS provides a signaling service, which can be used by multiple upper layers for several types of application. We describe this service here as an abstract set of capabilities. A later version of this framework could illustrate the use of these capabilities within a broader context (e.g. how NSIS signaling could be used within a complete set of message flows that signal a voice over IP call).

We can loosely define the boundary between NSIS and these upper layers from three views:
*) What basic control primitives are available at the interface;
*) What information is exchanged within these primitives;
*) What assumptions NSIS makes about operations carried out above the interface.

The set of control primitives required is quite small.
At the initiating (NI) end:
*) UL requests signaling for a new resource;
*) UL requests modification or removal of an existing resource.
*) UL receives progress indications (minimally, success or failure).
At the responding (NR) end:
*) Notification to UL that a resource has been set up.
At either end:
*) Notification to UL that something has changed about the available resource and other error conditions.
This description is in terms of a 'hard state' interface, without explicit refresh messages between upper layers and NSIS, although this is an implementation issue. In any case, NSIS implementations will need to be able to detect conditions when ULs fail without issuing explicit resource removal requests.

The information in the control primitives consists essentially of two parts. The first is the definition of the data flow for which the resource is being signaled. The format (e.g. socket id or packet fields or whatever) is an implementation issue; it has to be interpreted into a 'wire format' (as in section 4.5). Since NSIS could support both sender and receiver initiation, the flow definition must also state whether it is incoming or outgoing over a particular interface (this can be inferred when the initiator is colocated with the flow endpoint). The second part of the information exchanged is the service definition (e.g. QoS description in the case of a QoS request). This is opaque to NSIS, with the possible exception of identifying the resource type being signaled.

We have a basic design goal not to duplicate functionality that is already present in (or most naturally part of) existing signaling protocols which could be used by the upper layers. Therefore NSIS

(implicitly) assumes that certain procedures are carried out 'externally'. The main aspects of this are:
*) Negotiation of service configuration (e.g. discovering what services are available to be requested);
*) Agreement to use NSIS for signaling, and coordination of which end will be the initiator;
*) (Potentially) discovery of the NSIS peer to be signaled with, especially if this is not directly on the data path. See also the security discussion in section 6.
Actually providing these functions might require enhancements to these other protocols. These are still to be identified.

4.3 Protocol Structure

4.3.1 Internal Layering

We can model the NSIS protocol as consisting of three layers, as shown in Figure 2. This is initially just a way of grouping associated functionality, and does not mean that all these layers could necessarily operate or even be implemented independently.

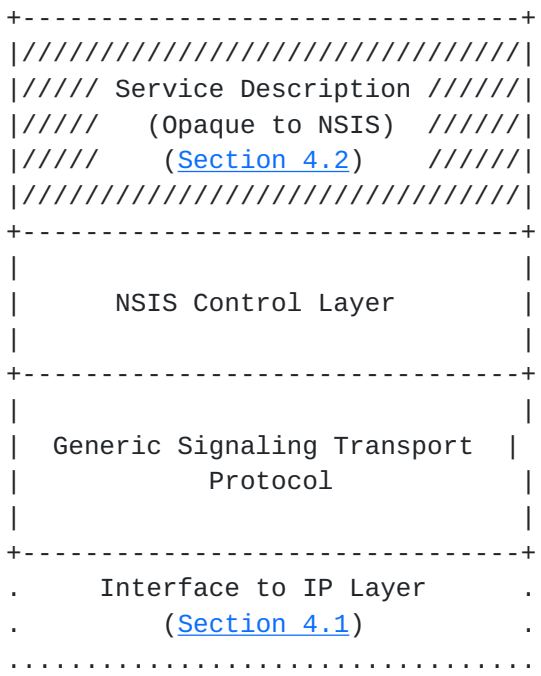

Figure 2: NSIS Layer Structure

The lower layer interface (to IP) has been described in section 4.1. The service description information is essentially the same as provided by the upper layers, as described in section 4.2. It isn't clear if the service description can be independent of the lower parts of the protocol or whether different descriptions would be

valid at different stages of protocol operation. This depends on the particular service, and therefore to make NSIS service independent we must allow that the service description part may be explicitly dependent on the 'NSIS' fields which lie below. This is similar to the ALSP/CSTP coupling described in [7].

The distinction between the 'NSIS layer' and the 'Generic Signaling' layer is not functionally clear cut, but one of convenience. In outline:
*) The 'generic' layer provides (at most) functionality which might be available from existing protocols, such as SCTP [8] or IPSec [9]. An extreme case could be the binding update messages of mobility signaling (section 5.3.4).
*) The 'NSIS' layer provides (at least) functionality which is somehow specific to path-directed signaling.

Functionality reasonable to re-use from existing signaling protocols might include reliability and re-ordering protection, dead peer detection (keepalive), multihoming support, payload multiplexing (piggybacking), and security services, such as establish a security context and carrying out key exchange.

Functionality which would probably have to be in the NSIS layer would include flow and reservation identification, some error handling, demultiplexing between different resource types, as well as the basic NSIS messages. More details on the messages are in section 4.3.2 and the identifier aspects in section 4.5.

The choice of using functionality from an existing protocol or re-specifying it as part of NSIS is for further analysis. It probably depends on the function in question, and in the end might be left flexible to allow optimization to local circumstances. (For example, Diameter allows the use of IPSec for security services, but also includes its own CMS application as an alternative.) Whichever approach is taken, the combination of NSIS and supporting transport protocol must provide a uniform protocol capability to the service layer.

4.3.2 Protocol Messages

The NSIS specific part protocol will include a set of messages to carry out particular operations along the signaling path. Initial work for RSVP concentrated on the particular case of QoS reservation signaling, although in principle, the necessary basic messages could depend on the resource type NSIS is being used for. However, the implication of the analysis in [7] is that this message set generalizes to a wide variety of signaling scenarios, and so we use it as a starting point. A very similar set was generated in [10].

Name	Direction	Semantics
Request	I-->R	Create a new reservation for a flow
Modify	I-->R (&R-->I?)	Modify an existing reservation
Release	I-->R & R-->I	Delete (tear down) an existing reservation
Accept/ Reject	R-->I	Confirm (possibly modified?) or reject a reservation request
Notify	I-->R & R-->I	Report an event detected within the network (e.g. congestion condition or end of condition)
Refresh	I-->R	State management (see section 4.4)

Note that the 'direction' column in this table only indicates the 'orientation' of the message. The messages can be originated and absorbed at NF nodes as well as the NI or NR; an example might be NFs at the edge of a domain exchanging NSIS messages to set up resources for a flow across a it.

Note the working assumption that responder as well as the initiator can release a reservation (comparable to rejecting it in the first place). It is left open if the responder can modify a reservation, during or after setup. This seems mainly a matter of assumptions about authorization, and the possibilities might depend on resource type specifics.

The table also explicitly includes a refresh message. This does nothing to a reservation except extend its lifetime, and is one possible state management mechanism for NSIS. This is considered in more detail in section 4.4.

4.4 State Management

The prime purpose of NSIS is to manage state information along the path taken by a data flow. There two critical issues to be considered in building a robust protocol to handle this problem:
*) The protocol must be scalable. It should minimize the state storage demands that it makes on intermediate nodes; in particular,

storage of state per 'micro' flow is likely to be impossible except at the very edge of the network.
*) The protocol must be robust against failure and other conditions, which imply that the stored state has to be moved or removed.

The total amount of state that has to be stored depends both on NSIS and on the resource type it is being used to signal for. The resource type might require per flow or lower granularity state; examples of each for the case of QoS would be IntServ or RMD (per 'class' state) respectively. The NSIS protocol should not overburden an application that was otherwise lightweight in state requirement. However, depending on design details, it might require storage of per-flow state including reverse path peer addressing, simply for sending NSIS messages themselves.

There are several robustness problems, which roughly align with the 'layers' of the NSIS protocols of Figure 2, that can be handled by the soft state principle. (Independence of these layers therefore implies the danger of duplication of functionality.) This relies on periodic refresh of the state information with the current context, relying on invalid state being timed out. Soft state can be used either as the primary mechanism to handle the problem, or sometimes as a backup to some other approach.

*) At the lowest level, soft state can be used to detect dead NSIS peers - loss of several periodic messages implies termination of the signaling. (The same inference can be made e.g. if failure is detected at the link layer.) The assumption is then that the corresponding reservation should be automatically deleted, and the deletion propagated along the remainder of the path.

*) At the next level, in the event of a routing change (for example caused by network changes or end host mobility), reservation state should be removed from the old path and added to the new one. This will be handled automatically by periodic messaging, provided that the entities on the new path accept a Refresh message to install a new reservation. (A partial alternative is to have a routing-aware NSIS implementation, if the route change takes place at an NSIS-aware node.)

*) At the highest level, a particular resource type might have timing limits associated with a particular reservation (e.g. credit limited network access). Periodic re-authorized requests can be used as part of the time control.

All of these can be handled with a single soft state mechanism, although it may be hard to choose a single refresh interval and message loss threshold appropriate for all of them. Even where

alternative approaches are possible, for example using knowledge of the fact that a routing change has occurred to trigger an explicit NSIS release message, it seems that a soft state mechanism is always necessary as a backup.

4.5 Identity Elements

NSIS will carry certain identifiers within the NSIS layer. The most significant identifier needs seem to be the following.

4.5.1 Flow Identification

The flow identification is a method of identifying a flow in a unique way. All packets and/or messages that are associated with the same flow will be identified by the same flow identifier. In principle, it could be a combination of the following information (note that this is not an exclusive list of information that could be used for flow identification):

*) source IP address;
*) destination IP address;
*) protocol identifier and higher layer (port) addressing;
*) flow label (typical for IPv6);
*) SPI field for IPSec encapsulated traffic;
*) DSCP/TOS field

We've assumed here that the flow identification is not hidden within the service definition, but is explicit as part of the basic NSIS protocol. The justification for this is that it might be valuable to be able to do NSIS processing even at a node which was unaware of the specific resource type and service definitions in question; this would be a case of an NSIS forwarder with no interface to any resource management function. An example scenario would be NSIS messages passing through an addressing boundary where the flow identification had to be re-written.

The very flexibility possible in flow classification is a possible source of difficulties: when wildcards or ranges are included, it is probably unreasonable to assume a standard classification capability in routers; on the other hand, negotiating this capability would be a significant protocol complexity.

4.5.2 Reservation Identification

There are several circumstances where it is important to be able to refer to a reservation independently of whatever other information is associated with it. The prime example is a mobility-induced address change (handover) which required the flow identifier associated with a reservation to be rewritten without installing a totally new

reservation (see section 5.3.1 for some security and scoping implications of this use). The same capability could also be used to simplify refresh or release messages in some circumstances, and might be useful within the protocol to resolve reservation collisions (where both sender and receiver initiate for the same flow).

A reservation identifier performs these roles. It is open how the reservation identifier space should be defined and managed, and what the scope of the identifier should be (only peer-peer, or end-end, when interpreted in conjunction with some of the addressing information). Some of the necessary identifier functions, especially to do with local operation of NSIS, may also be provided by lower layer signaling transport protocols.

4.5.3 Resource Type Identification

Since NSIS can be used to support several uses, there is a need to identify which resource type a particular NSIS invocation is being used to signal for, and this needs to be done outside the (opaque) service description:
*) processing incoming request messages at a responder - the NSIS layer should be able to demultiplex these towards the appropriate upper layer;
*) processing general NSIS messages at an NSIS aware intermediate node - if the node does not handle the specific resource type, it should be able to make a forwarding decision without having to parse the service description.

Resource type identifiers would probably require an IANA registry.

5. NSIS and other Functions and Protocols

5.1 Resource Management and Network Provisioning

It is a requirement for the NSIS protocol to be independent of resource allocation and management techniques that may be used in the network. As such, we need to define the interaction between NSIS and what we will call the Resource Management Function (RMF). The RMF is responsible for all network provisioning and resource allocation functions.

In its resource provisioning role, the RMF can act as a client towards the NSIS protocol, as a particular "application" triggering an NI for resources in the network. This situation is depicted in Figure 3 and Figure 4.

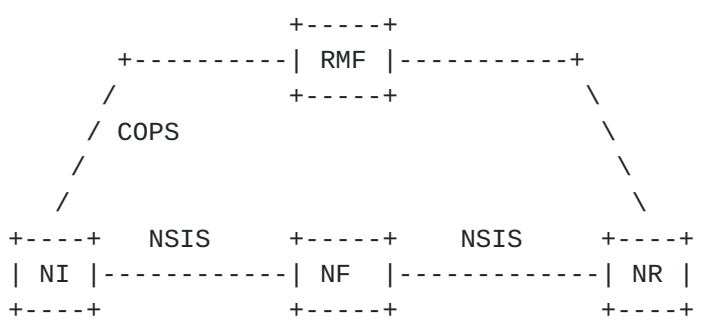

Figure 3: Centralized RMF as a client to NSIS

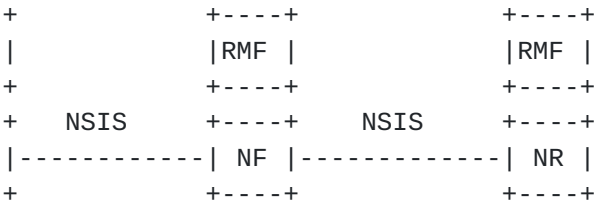

Figure 4: Distributed RMF as a client to NSIS

When the RMF is distributed in the network, a protocol for communication with the NI, NF, NR may not be required. In this case the RMF is providing traffic classification and conditioning functions; an example of such functionality is described in [11].

Conversely, the RMF can be a server to an NI, NF or NR controlling a complete domain. In the centralized case, it would be natural to formalize the relation between the nodes containing NEs and the central RMF as a Service Level Agreement (SLA). In order to shield the NE from (resource specific) SLA aspects, we would model the interaction as being via some kind of local 'proxy' the RMF. This situation is depicted schematically in Figure 5. Figure 6 shows the corresponding distributed case. Note that the functional split between the NE and RMF is the same in each case; in other words the same NSIS functionality supports both scenarios.

In case of centralized RMF, the SLA or its technical part, the Service Level Specification (SLS) [12] specifies the resource guarantees that the RMF needs to provide. These guarantees apply between one or more ingress and egress points of the network. The SLS also specifies the availability and reliability of the service. In the case of QoS signaling, it may refer to a bandwidth service with certain performance guarantees regarding delay, jitter or packet loss.

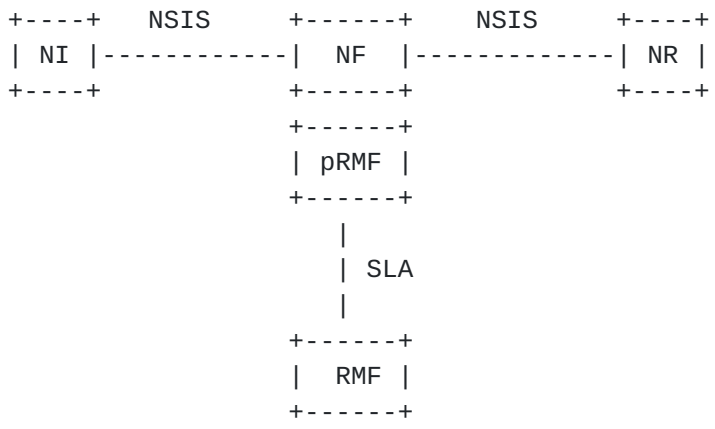

Figure 5: Centralized RMF as a server to NSIS

Figure 6: Distributed RMF as a server to NSIS

The decoupling of NSIS signaling and network management by means of an SLS has some attractive properties:
- It allows a Network Provider to easily share the use of its infrastructure between several Service Providers using NSIS signaling to provide their service.
- It allows a clear separation between resource provisioning and management and reservation signaling and admission control.
- It relieves the NF from several tasks, making it potentially more scalable in the core of the network.

The resource management system can perform either per-flow or per-class admission control decisions based on the requested QoS information and on the reservation state it keeps regarding active flows (or classes). Keeping per-flow state may be required for policing, accounting/billing and explicit reservation teardown. Per-flow based functions can be mandatory in some parts of the network, e.g., end host to first hop router, or at the edge of the network or at the boundary of a network domain. Conveniently, this is also where the processing needed to maintain per-flow state will remain manageable. In the core, this approach may not scale very well and per-class state may be used as an alternative that is very scalable and allows for a lightweight processing of signaling messages. With per-class state, however, we lose the ability to directly notify the NE in case of unsolicited network events because the affected flows

cannot be identified. Instead, the situation needs to be detected from the response to a refresh message which in turn mandates the use of soft-state with separate messages or message structure for requests and refreshes.

The RMF can execute its network provisioning functions according to its internal policies. In the easiest case, it may run an overprovisioned network with only monitoring capabilities in order to follow up on the delivered performance. In more complex scenarios, it may use a whole array of network optimization tools in order to deliver and maintain service quality according to the SLS.

5.2 IP Routing

Several situations may occur when routing diverges from standard layer 3 routing. These are summarized in the sections below.

5.2.1 Load Sharing

Load sharing or load balancing is a network optimization technique that exploits the existence of multiple paths to the same destination in order to obtain benefits in terms of protection, resource efficiency or network stability. The significance of load sharing in the context of NSIS is that, if the load sharing mechanism in use will forward packets on any basis other than source and destination address, routing of NSIS messages using end-to-end addressing does not guarantee that the messages will follow the data path. In this section, we briefly survey what standard methods have been used for load sharing within standard routing protocols.

In OSPF, load balancing can be used between equal cost paths [13] or unequal cost paths. An example of the latter approach is Optimized Multi Path (OMP). OMP discovers multiple paths, not necessarily equal cost paths, to any destinations in the network, but based on the load reported from a particular path, it determines which fraction of the traffic to direct to the given path. Incoming packets are subject to a (source, destination address) hash computation, and effective load sharing is accomplished by means of adjusting the hash thresholds.

BGP [14][15] advertises the routes chosen by the BGP decision process to other BGP speakers. In the basic specification, routes with the same Network Layer reachability information (NLRI) as previously advertised routes implicitly replace the original advertisement, which means that multiple paths for the same prefix cannot exist. Recently, however, a new mechanism was defined that will allow the advertisement of multiple paths for the same prefix without the new paths implicitly replacing any previous ones [16]. The essence of the

mechanism is that each path is identified by an arbitrary identifier in addition to its prefix.

The distribution of traffic over the available path may be done per destination, per message in a round-robin fashion or with a predefined hashing function. The determination of the hashing image may take into account the source/destination IP address, QoS information such as the DSCP or protocol ID. When the routing decision is no longer based on the destination address only, however, there is a risk that data plane messages and control plane messages will not follow the same route.

5.2.2 QoS Routing

The are several proposals for the introduction of QoS awareness in the routing protocols. All of these essentially lead to the existence of multiple paths (with different QoS) towards the same destination. As such, they also contain an inherent risk for a divergence between control plane and data plane, similar to the load sharing case.

For intra-domain traffic, the difference in routing may result from a QoS-aware traffic engineering scheme, that e.g. maps incoming traffic to LSPs based on multi-field classification. In BGP, several techniques for including QoS information in the routing decision are currently proposed. A first proposal is based on a newly defined BGP-4 attribute, the QoS_NLRI attribute [17]. The QoS_NLRI attribute is an optional transitive attribute that can be used to advertise a QoS route to a peer or to provide QoS information in along with the Network Layer Reachability Information (NLRI) in a single BGP update. A second proposal is based on controlled redistribution of AS routes [18]. It defines a new extended community (the redistribution extended community) that allows a router to influence how a specific route should be redistributed towards a specified set of eBGP speakers. The types of redistribution communities may result in a specific route not being announced to a specified set of eBGP speakers, that it should not be exported or that the route should be prepended n times.

5.2.3 Route pinning

Route pinning refers to the independence of the path taken by certain data packets from reachability changes caused by routing updates from an Interior Gateway Protocol (OSPF, IS-IS) or an Exterior Gateway Protocol (BGP). This independence may for instance be caused by the configuration of static LSPs or by the establishment of explicitly routed LSPs by means of a signaling protocol (RSVP-TE or CR-LDP). If the NSIS signaling messages follow standard Layer 3 routing, this may cause a divergence between control plane and data plane. If

reservations are made on the control plane, this may result in sending data along an unreserved path while maintaining a reservation on a path that is not used.

5.2.4 Route changes

In this section, we will explore the expected interworking between a signaling for resource BGP routing updates, although the same applies for any source of routing updates. The normal operation of the NSIS protocol will lead to the situation depicted in Figure 7, where the reserved resources match the data path.

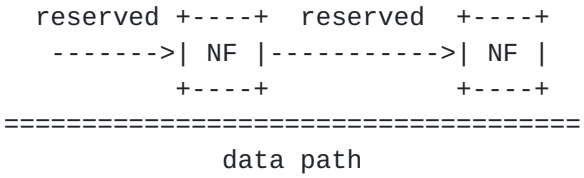

Figure 7: Normal NSIS protocol operation

A route change (triggered by a BGP routing update for instance) can occur while such a reservation is in place. In case of RSVP, the route change will be installed immediately and any data that is sent will be forwarded on the new path. This situation is depicted Figure 8.

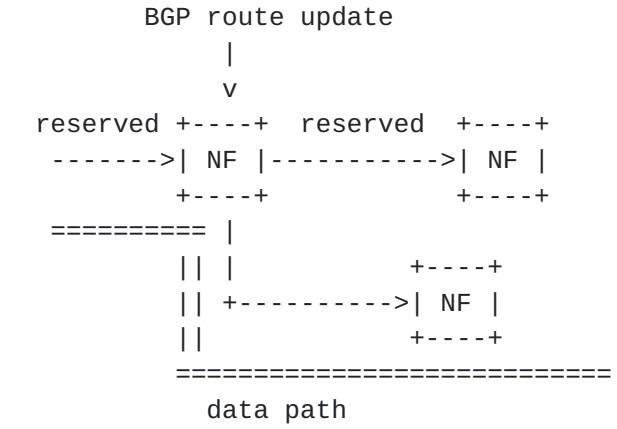

Figure 8: Route Change

Resource reservation on the new path will only be started once the next control message is routed along the new path. This means that there is a certain time interval during which resources are not reserved on (part of) the data path. To minimize this time interval several techniques could be considered. As an example, RSVP [19] has the concept of local repair, where the router may be triggered by a route change. In that case the RSVP node can start sending PATH

messages directly after the route has been changed. Note that this option may not be available for NSIS if no per-flow state is kept in the NF.

It is not guaranteed that the new path will be able to provide the same guarantees that were available on the old path. Therefore, in a more desirable scenario, the NF should wait until resources have been reserved on the new path before installing the route change. The route change procedure then consists of the following steps:
1. NF receives a route announcement,
2. Refresh messages are forwarded along the current path,
3. A copy of the refresh message is remarked as request and send along the new path that was announced,
4. When the NF has been acknowledged about the reservations on the new path the route will be installed and the traffic will flow along the new path.

Another example related to route changes is denoted as severe congestion and is explained in [20]. This solution adapts to a route change, when a route change creates a congestion on the new routed path.

5.3 Mobility Support

The interactions between mobility and resource signaling protocols have been quite extensively analyzed in recent years, primarily in the context of RSVP and Mobile IP interaction (e.g. [21]), but also in the context of other types of network (e.g. [22]). This analysis work has shown that some difficulties in the interactions are quite deep seated in the detailed design of these protocols; however, the problems and their possible solutions fall under five broad headings. The main issue is to limit the period after handovers during which the resource state has not been installed on the path, in particular the new part of the path.

We can use this work as the starting point for considering the framework aspects of a new resource signaling protocol like NSIS, which will need to interwork with mobility signaling, e.g., Mobile IP, or mobility paradigms using micromobility, or application layer approaches.

5.3.1 Addressing and Encapsulation

A mobility solution typically involves address reallocation on handover (unless a network supports per host routing) and may involve special packet formats (e.g. the routing header and Home Address option of MIPv6). Since NSIS may depend on end system addresses for forwarding signaling messages and defining flows

(section 4.5.1), the special implications of mobility for addressing need to be considered. Examples of possible approaches that could be used to solve the addressing and encapsulation problem are as follows:
*) Use a filter definition based on low level IP addresses (e.g. the Care of Address) and other 'standard' fields in the IP header. This makes least demands on the packet classification engines within the network. However, it means that even on a part of the flow path which is unchanged, the reservation will need to be modified to reflect the changed flow identification (see section 5.3.3).
*) Use a flow definition that does not change (e.g. based on Home Address); this is the approach assumed in [23]. This simplifies the problem of reservation update, at the likely cost of considerably complicating the flow identification requirements.

In the first approach, to prevent double reservation, NSIS nodes need to be able to recognize that a reservation with the new flow identifier is to be correlated with an existing one. The reservation identifier (section 4.5.2) was introduced for exactly this purpose. Note that this would require the reservation identifier to have (secure) end to end significance. (An additional optimization here would be use a local mobility management scheme to localize the visibility of the address change.)

The feasibility and performance of this approach needs to be assessed, including a detailed analysis of the signaling scenarios after a handover. However, given the high impact of requiring more sophisticated packet classifiers, initially it still seems more plausible than the second approach. This implies that the NSIS initiator should define flows in terms of real (care of) addresses rather than virtual (home) addresses. Thus, it would have detailed access to lower layer interface configuration (cf. section 4.1), rather than operating as a pure application level daemon as is commonplace with current RSVP implementations.

5.3.2 Localized Path Repair

In any mobility approach, a handover will cause at least some changes in the path of upstream and downstream packets. NSIS needs to install new state on the new path, and remove it on the old. Provided that some NSIS node on the joined path - the crossover router - can recognize this situation (which again depends on reservation identification), state installation and teardown can be done locally between it and the mobile node. (This may have implications for which entities are allowed to generate which message types, see section 4.3.2). It seems that the basic NSIS framework already contains the fundamental components necessary for this.

A critical point here is the signaling that is used to discover the crossover router. This is a generalization of the problem of finding next-NSIS-hop nodes: it requires extending the new path over several hops until it intersects the old one. This is easy for uplink traffic (where the mobile is the sender), but much harder for downlink traffic without signaling via the correspondent. There is no reason for the crossover routers for uplink and downlink flows to be the same, even for the same correspondent. The problem is discussed further in [24].

5.3.3 Reservation Update on the Unchanged Path

On the path between the crossover router(s) and the correspondent, it is necessary to avoid, if possible, double reservations, but rather to update the reservation state to reflect new flow identification (if this is needed, which is the default assumption of section 5.3.1). Examples of approaches that could be used to solve this problem are the following:

*) Use a reservation state definition that does not change even if the flow definition changes (see Section 4.5.2). In this case this problem is solved.

*) Use signaling all the way to the correspondent node (receiver end host), accepting the additional latency that this might impose.

*) Use an NSIS-capable crossover router that manages this reservation update autonomously (more efficiently than the end nodes), with similar considerations to the local path repair case.

5.3.4 Interaction with Mobility Signaling

In existing work on mobility protocol and resource signaling protocol interactions, several framework proposals describing the protocol interactions have been made. Usually they have taken existing protocols (Mobile IP and RSVP respectively) as the starting point; it should be noted that an NSIS protocol might operate in quite a different way. In this section, we provide an overview of how these proposals would be reflected in framework of NSIS. The mobility aspects are described using Mobile IP terminology, but are generally applicable to other network layer mobility solutions. The purpose of this overview is not to select or priorities any particular approach, but simply to point out how they would fit into our framework and point out any major issues with them.

We can consider that two signaling processes are active: mobility signaling (e.g. Binding updates or local micromobility signals) and NSIS. The discussion so far considered how NSIS should operate. There is still a question of how the interactions between the NSIS and mobility signaling should be considered.

The basic case of totally independent specification and implementation seems likely to lead to ambiguities and even interoperability problems (see [23]). At least, the addressing and encapsulation issues for mobility solutions that use virtual links or their equivalents need to be specified in an implementation-neutral way.

A type of 'loose' integration is to have independent protocol definitions, but to define how they trigger each other - in particular, how the mobility protocol triggers NSIS to send refresh/modify/tear messages. A pair of implementations could use these triggers to improve performance, primarily reducing latency. (Existing RSVP modification consider the closer interaction of making the RSVP implementation mobility-routing aware, e.g. so it is able to localize refresh signaling; this would be a self contained aspect of NSIS.) This information could be developed for NSIS by analyzing message flows for various mobility signaling scenarios as was done in [21].

An even tighter level of integration is to consider a single protocol carrying both mobility and resource information. Logically, there are two cases:
1. Carry mobility routing information (a 'mobility object') in the resource messages, as is done in [23]. (The prime purpose in this approach is to enable crossover router discovery.)
2. Carry resource signaling in the mobility messages, typically as a new extension header. This was proposed in [25] and followed up in [26]; [27] also anticipates this approach. In our framework, we could consider this a special case of NSIS layering, with the mobility protocol playing the role of the signaling transport (as in 4.3.1). The usefulness of this class of approach depends on a tradeoff between specification simplicity and performance. Simulation work is under way to compare the performance of the two approaches in the case of RSVP and micromobility protocols.

Other modes of interaction might also be possible. The critical point with all these models is that the general solutions developed by NSIS should not depend fundamentally on the choice of any particular mobility protocol. Especially if it has interdomain scope, tight integration would have major deployment issues; loose integration could require NSIS implementations to hook into multiple different mobility protocols. Therefore, any integrated solution should be considered out of scope of initial NSIS development, and even in the long term is probably only applicable if it can be localized within a particular part of the network.

5.3.5 Interaction with Fast Handoff Support Protocols

In the context of mobility between different access routers, it is common to consider performance optimizations in two areas: selection of the optimal access router to handover to, and transfer of state information between the access routers to avoid having to regenerate it in the new access router after handover. The seamoby working group is developing solutions for these protocols for pure IP based networks (CARD and CT respectively); other networks, which use NSIS for resource signaling within the network, may use different types of solution.

In this section, we consider how NSIS should interact with these functions, however they are implemented. Detailed solutions are not proposed, but the way in which interaction these functions is seen within the NSIS framework is described. NSIS should be able to operate independently of these protocols. However, significant performance gains could be achieved if they could be made to cooperate. In addition, the resource signaling aspects of these protocols could profitably use a common set of resource types and definitions with NSIS to avoid a proliferation of incompatible service models (also since at any given node, these protocols will probably interface to common resource management functions).

The question arises, what the mode of interaction should be: independent operation, NSIS triggering access router discovery and state transfer, or vice versa. The questions for the two cases seem to be independent.

For access router discovery, a typical model of operation is that the mobile carries out an information gathering exercise about a range of capabilities. In addition, where those capabilities relate purely to the AR and mobile, there is no role for NSIS (its special functionality is not relevant). However, considering resource aspects, one aspect of the AR 'capability' is resource availability on the path between it and the correspondent, and NSIS should be able to fulfill this part. Indeed, this is effectively precisely the application considered in [26], where it is a sort of special case of resource signaling during handover.

Therefore, a possible model of access router discovery/NSIS relationship is that some entity in a candidate AR triggers NSIS using resource and reservation information (including reservation id) from the current AR to find out about what would be available on the new path. Note that this should be a query rather than an actual reservation; this semantic could be included either in the service definition or NSIS itself.

The case of state transfer is more complex. There are two obvious options, corresponding to whether one transfer just resource state or NSIS state as well:

1. "State transfer triggering NSIS": A state transfer process passes the 'raw' resource state to the new AR. This triggers a new instance of NSIS to request that resource.
2. "NSIS using state transfer": NSIS transfers its own state information from the old to the new AR. It can then carry out the same update signaling as though it was a single 'virtual AR' which had just had a topology change towards the correspondent. (This is essentially the conceptual model of [21].)

The first model is simpler, and maybe more in line with the basic state transfer expectation; however, it seems hard to avoid double reservations since the two NSIS protocol instances are not coordinated. Therefore, the second model seems more appropriate. An advantage of the 'virtual AR' model is that it ensures that the impact of the interaction is limited to the NSIS instances at ARs themselves, since the rest of the network must be able to handle a topology change anyway.

Note that there is an open issue of who is responsible between the mobile and AR to decide that the state transfer procedures have not happened for whatever reason - e.g. because they were not even implemented - and take recovery action to have the mobile refresh reservations promptly. It appears this has to be an NSIS responsibility in the AR, and probably requires a custom notification message for this circumstance.

5.4 Existing Resource Signaling Protocols

It is hoped that an NSIS protocol could eventually achieve widespread use for resource signaling. However, it is bound to have to inter-operate with existing resource signaling protocols at least during transition and possibly long term. The prime example here is RSVP, although other proprietary or domain specific protocols (e.g. bandwidth broker related) may also be considered. A related issue is that NSIS will be only one part of a resource control solution: it will always need to interwork with other resource-related protocols (e.g. COPS).

Analyzing the constraints on NSIS that come from these requirements is hard before further refinement of the framework has been carried out and critical assumptions pinned down. However, we can identify various modes of interoperation, and the attributes of the framework that will make them easy.

Firstly, we should allow for NSIS to be used over a 'long range', in conjunction with a different protocol locally (e.g. intra-domain); or, the two roles could be reversed. This is actually very similar to the case of use of NSIS layered over itself (section 5.5). In the case where the 'inter-layer' interaction is mediated via resource management, the same should approach should work with non-NSIS protocols. What needs to be validated here is whether NSIS layering requires the exchange of NSIS specific information between the layers.

A second issue is that NSIS should be able to be deployed within an environment without radical changes to supporting resource (or AAA) related protocols. The main issue here is that NSIS should be flexible in its ability to support different service definitions (and possibly flow classifications). This is already one of the main goals of the framework presented here.

The final point is that it should be possible to use NSIS over one network region, concatenated with another protocol over an adjacent region. The main issue here, apart from the flexible service and flow capabilities already mentioned, is that NSIS should be adaptable in what signaling paths (e.g. to interwork with both on- and off-path solutions), and in initiation paradigms (e.g. to interwork with sender and receiver initiated solutions).

5.5 Multi-Level NSIS Signaling

This section describes a way of separating the NSIS signaling protocol into more than one hierarchical level. In this section three levels of hierarchy are considered (see Figure 9); however, the approach is quite general to more (or fewer) levels: the important issue is the use of NSIS at more than one level at all.

The lowest hierarchical level ("level 1") provides basic resource management functionality related to scalable, simple and fast soft state maintenance and to transport functions, such as reliable delivery of signaling messages, congestion control notification and load sharing adaptation. Soft state that is maintained by this level is usually per traffic class based.

The second hierarchical level ("level 2") is more complex than level 1 as regards soft state maintenance. Soft state maintained by this hierarchical level is usually per flow. Note that this level, like level 1, also supports transport functions. When an NSIS edge-to-edge multi-domain protocol is used, level 2 stretches beyond domain boundaries and is applied on all the edges of the domains that are included in the multidomain region.

The third hierarchical level ("level 3") includes a set of upper-level signaling functions that are specific to particular signaling applications. Such functions could, for example, be security, policy, billing, etc.

As shown in Figure 9, the three hierarchical levels might be applied on different NSIS entities.

This three-level architecture for NSIS signaling can be provided by using:

* a single end-to-end NSIS protocol that supports all three hierarchical levels

* two independent NSIS protocols: Level 3 is supported by an end-to-end NSIS protocol, and levels 1 and 2 are supported by another edge-to-edge NSIS protocol.

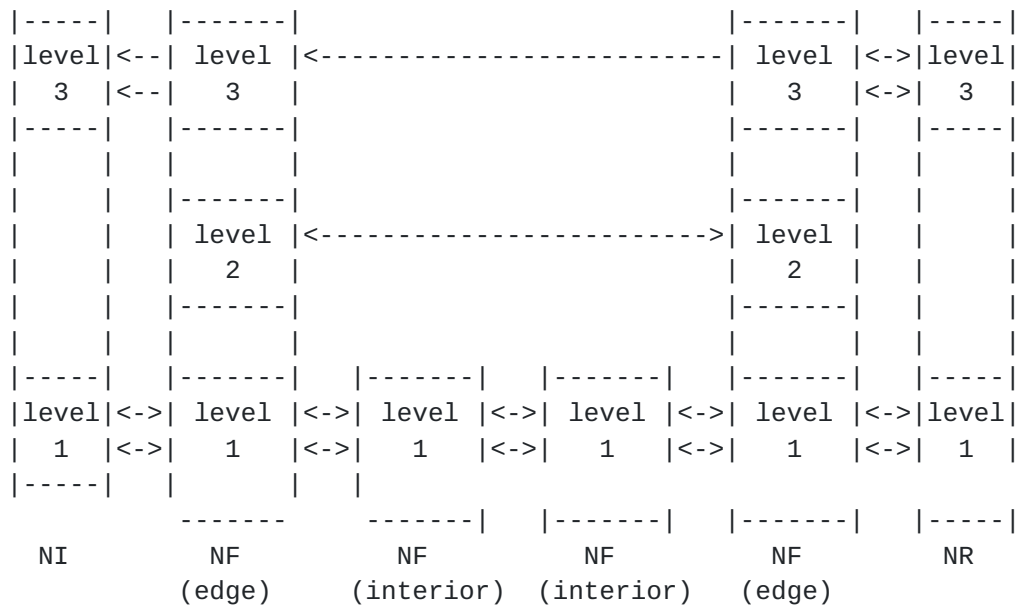

Figure 9: Three level architecture for NSIS signaling

* NI (NSIS Initiator): can be an end-host or a proxy and can process and use the "level 1" and "level 3" protocol components

* NR (NSIS Responder): can be an end-host or a proxy and can process and use the "level 1" and "level 3" protocol components

* NF (NSIS Forwarder) (edge): can be a Diffserv edge, MPLS edge, etc. It can process and use the "level 3", "level 2" and "level 1" protocol components. Usually,

"level 2" provides an interworking between "level 1" and "level 3" protocol components.

* NF (interior): can be any router within a domain. It can process and use only the "level 1" protocol component. The "level 3" and "level 2" protocol components are not processed (used or checked);

The hierarchical level separation can be provided by supporting a hierarchical object structure. In other words, the NSIS protocol objects should be structured and positioned within the NSIS messages in a hierarchical way, i.e., first the "level 1" objects, then the "level 2" objects and finally the "level 3" objects.

6. Security and AAA Considerations

A framework is meant to create boundaries for a later protocol and to describe the interaction between the protocol and its environment. Security issues usually turn out to have impacts in the interaction of these protocols and must therefore be appropriately addressed in such a framework. This section describes these general security issues, and in particular considers the interactions between NSIS and authentication, authorization and accounting. Together with authentication the protection of the signaling messages is addressed - namely replay and integrity protection.

An initial analysis of the major security threats that apply in the typical of scenario where NSIS is expected to be used is given in [4]; these threats are described at the overall scenario level, in terms of the impact on users and networks. However, in any given scenario, NSIS will be just one protocol or component of the overall solution. Ultimately, the framework will need to what aspects of these threats need to be handled by NSIS compared to the other components. Currently, we can only make initial scoping assumptions of this sort.

6.1 Authentication

Authentication (and key establishment) for a signaling protocol should be seen as a two-phase process. The first-phase is usually more performance intensive because of a larger number of roundtrips, denial of service protection, cross-realm handling, interaction with other protocols and the likely larger cryptographic computation associated with it. As stated in section 4.3, this functionality could be provided externally to NSIS, e.g. by reusing a standard transport protocol which already included this functionality. At the end of this phase it should be possible to create or derive security associations that are usable for the protection of the NSIS signaling

messages themselves. The functionality required here relates to (data origin) authentication (including integrity and replay protection) of individual signaling messages. Key establishment, rekeying, synchronization issues are issue that may be addressed here depending on the specific method. In any case the protection applied to each signaling message must be fast and efficient.

When using cryptography to protect signaling messages, it is obvious that a node must be able to select the appropriate security association in order to be able to apply signaling message protection. This should just be a general point about endpoint identity issues. Hence the identity identifier must be available to the transmitting node. Regarding identities there is a need to support different identity types to enable the flexible usage of several signaling initiators and receivers. Supporting static configuration and dynamic learning of these identities should be provided.

6.2 Authorization

Authorization information can be seen in an abstract form as "Can the resource requestor be trusted to pay for the reservation?". This abstraction is supported by the fact that reservations require some form of incentive to use some 'default' resource (or vice versa - penalty for not reserving too many resources). In general, the semantics of the authorisation will depend on the type of resource (QoS, firewall configuration etc.) that NSIS is being used to signal for. The implication of this is that NSIS will not directly make authorisation decisions; instead, the authorisation information must be fed into the resource management function (section 5.1) which actually decides the allocation (or rejection) of the request.

Some negotiation needs to take place to determine which node will take responsibility for authorising a resource request, the implication being that the same node will ultimately be accounted to for it. Such a negotiation needs to be flexible enough to support most currently deployed schemes (e.g. reverse charging, etc.) while keeping efficiency and simplicity in mind. This negotiation might be executed before starting resource signaling (assumed in section 4.2), although it could also be part of the NSIS signaling messages (as in some proposals dealing with charging and RSVP). Since information needs to be sent to the networks, some information needs to be included to provide the network with the necessary information to start the authorisation process. Hence fully opaque objects might not always be the proper choice.

It is not clear if 'initiation' of a reservation is related to willingness to accept authorisation responsibility. (Current

practices tend to assume that flow originators are responsible.) In any case, it seems unlikely that a domain will make a cost-incurring request of a peer domain without already having received a matching request from the peer in the other direction - in other words, requests must propagate between domains in the same direction as authorisation responsibility. If this argument is correct, and if NSIS initiation and authorisation responsibility are decoupled, it must be possible for the authorisation responsibility to propagate both in the direction initiator->responder and vice versa. Also, if both [flow] sender and receiver initiation are possible, service descriptions must include information about the authorisation policy to be applied, which must be imposed consistently along the whole path. These issues should be analyzed to determine if 1, 2 or 4 alternative scenarios are possible and realistic.

A second question is that of which entities actually authorise which. One end user must ultimately get authorisation for the request (this may or may not be assumed to be the NSIS initiator, see below). There are then two possible models for how this authorisation is done throughout the path.

The first model assumes that each network along the path is able to authenticate and authorise the user directly. The implication for a signaling protocol is that the user credentials cannot be removed after the first hop and have to be further included in the message when forwarded to other networks. Every node along the path is then able to verify the user and to provide policy based admission control.

The second model assumes that the user credentials are removed at the first hop. The first network knows the user identity requesting the resources but does not include this information further along the path. The first network can therefore be seen as acting on behalf of the originator to take responsibility to enable further reservations to be done along the path i.e. in particular to the next network only. This procedure is then applied in a hop-by-hop basis.

Note that both models are independent on whether a traditional subscription based approach or an alternative means of payment (such as pre-pay on on-line charging by the visited network) is used. These issues only have an impact for the transmission of accounting records and for a requirement to execute an online verification whether a user still has sufficient credits/funds; therefore, these details do not affect NSIS operation.

6.3 Accounting

It is obvious that accounting/charging is an important part for the success and the acceptance of a resource signaling protocol. Most of the thinking in this area is derived from the specific case of signaling for QoS; however, we make an initial working assumption that the same paradigms should apply to signaling for any type of resource for which accounting is necessary. We can only refer to QoS as an example. We make the general assumption here that accounting records are generated by the resource management function based entirely on traffic measurements and processed in accordance with the authorisation information that was used in deciding to grant the request in the first place.

Therefore, NSIS plays no further part in this activity; the accounting records are transmitted using the AAA infrastructure, and charging and billing for the overall service is carried out at some higher layer. This would include feedback to applications (and users) about total session cost (of which the network resource cost might be only a part). An open issue is whether a query (without actually making a reservation) to the network should also generate a chargeable event; this could be considered as an aspect of the service definition.

6.4 End-to-End vs. Peer-Session Protection

It is reasonable to assume that peer-session security (with chain-of-trust) is used for most signaling environments relevant to NSIS. Especially the separation of signaling into different network parts (intra-domain within the access network, end-node to access network, intra-domain, and so on) and new proposals regarding mobility and proxy support show that the traditionally end-to-end signaling nature is not applicable in every environment (or possibly only in a minor number of environments). End-to-end security in a signaling protocol is actually problematic for two reasons:

a) Even if the messages use the address of the end-host (to support routing) if in path signaling is used then still the messages have to be interpreted and modified along the path.

b) The only property that can be achieved by using end-to-end security is that one end-host can be assured that the other end-host included some parameters (possibly resource parameters) that have not been modified along the path. Nodes along the path usually do not have the possibility to cryptographically verify the protected message parts. If the two end-points negotiate which side has to pay for the reservation (or possibly how much and other parameters) within the signaling protocol then there is a need to protect this

information. This leads to the question which protocols are executed before the signaling message exchange starts. If resource parameters and payment/charging related information are already exchanged beforehand as part of a separate protocol (possibly SIP) then there is little need to protect (and possibly retransmit) this information at the NSIS level basis. In most cases an opaque token to link the different protocols may be sufficient.

References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2 Brunner, M., "Requirements for QoS Signaling Protocols", draft-ietf-nsis-req-02.txt (work in progress), May 2002

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

4 Tschofenig, H., "NSIS Threats", draft-tschofenig-nsis-threats-00.txt (work in progress), May 2002

5 Katz, D., "IP Router Alert Option", RFC 2113, February 1997

6 Partridge, C., A. Jackson, "IPv6 Router Alert Option", RFC 2711, October 1999

7 Braden, R., "A Two-Level Architecture for Internet Signaling", draft-braden-2level-signal-arch-00.txt (work in progress), November 2001

8 Stewart, R. et al., "Stream Control Transmission Protocol", RFC 2960, October 2000

9 Kent, S., R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998

10 Westberg, L., et al., "Framework for Edge-to-Edge NSIS Signaling", draft-westberg-nsis-edge-edge-framework-00.txt (work in progress), May 2002

11 Blake, S., et al., "An Architecture for Differentiated Services", RFC2475, December 1998

12 Goderis, D., et al. "Service Level Specification Semantics and Parameters", draft-tequila-sls-02.txt (work in progress), February 2002

13 Apostolopoulos, G., et al., "QoS Routing Mechanisms and OSPF Extensions", RFC 2676, August 1999

14 Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995

15 Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", draft-ietf-idr-bgp4-17.txt (work in progress), January 2002

16 Walton, D., D. Cook, A. Retana and J. Scudder, "Advertisement of Multiple Paths in BGP", draft-walton-bgp-add-paths-00.txt (work in progress), May 2002

17 Cristallo, G., C. Jacquenet, "Providing Quality-of-Service Indication by the BGP-4 Protocol: the QoS_NLRI Attribute", draft-jacquenet-qos-nlri-04.txt (work in progress), March 2002

18 Bonaventure, O., S. De Cnodder, J. Haas, B. Quoitin and R. White, "Controlling the redistribution of BGP Routes", draft-bonaventure-bgp-redistribution-02.txt (work in progress), February 2002

19 Braden, R. et al., "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997

20 Westberg, L., M. Jacobsson, G. Karagiannis, S. Oosthoek, D. Partain, V. Rexhepi, R. Szabo, P. Wallentin, "Resource Management in Diffserv (RMD) Framework", draft-westberg-rmd-framework-01.txt (work in progress), February 2002

21 Thomas, M., "Analysis of Mobile IP and RSVP Interactions", draft-thomas-seamoby-rsvp-analysis-00.txt (work in progress), February 2001

22 Partain, D. et al., "Resource Reservation Issues in Cellular Radio Access Networks", draft-westberg-rmd-cellular-issues-01.txt (work in progress), June 2002

23 Shen, C. et al., "An Interoperation Framework for Using RSVP in Mobile IPv6 Networks", draft-shen-rsvp-mobileipv6-interop-00.txt (work in progress), July 2001

24 Manner, J., et al., "Localized RSVP", draft-manner-lrsvp-00.txt (work in progress), May 2002

25 Chaskar, H. and R. Koodli, "A Framework for QoS Support in Mobile IPv6", draft-chaskar-mobileip-qos-01.txt (work in progress), March 2001

26 Fu, X., et al, "QoS-Conditionalized Binding Update in Mobile IPv6", draft-tkn-nsis-qosbinding-mipv6-00.txt (work in progress), January 2002

27 Kan, Z., "Two-plane and Three-tier QoS Framework for Mobile IPv6 Networks", draft-kan-qos-framework-00.txt (work in progress), April 2002

Acknowledgments

The authors would like to thank Anders Bergsten, Maarten Buchli and Hannes Tschofenig for significant contributions in particular areas of this draft. In addition, the authors would like to acknowledge Marcus Brunner, Danny Goderis, Eleanor Hepworth, Cornelia Kappler, Hans De Neve, David Partain, Vlora Rexhepi, and Lars Westberg for insights and inputs during this and previous framework activities.

Author's Addresses

Ilya Freytsis
Cetacean Networks Inc.
100 Arboretum Drive
Portsmouth, NH 03801
USA
email: ifreytsis@cetacean.com

Robert Hancock
Roke Manor Research
Old Salisbury Lane
Romsey
Hampshire
SO51 0ZN
United Kingdom
email: robert.hancock@roke.co.uk

Georgios Karagiannis
Ericsson EuroLab Netherlands B.V.
Institutenweg 25
P.O.Box 645
7500 AP Enschede
The Netherlands
email: Georgios.Karagiannis@eln.ericsson.se

John Loughney
Nokia Research Center
11-13 Italahdenkatu
00180 Helsinki
Finland
email: john.loughney@nokia.com

Sven Van den Bosch
Alcatel
Francis Wellesplein 1
B-2018 Antwerpen
Belgium
email: sven.van_den_bosch@alcatel.be

Full Copyright Statement