Internet Engineering Task Force
Internet Draft

NSIS
H. Tschofenig
D. Kroeselberg
Siemens
Document:
draft-ietf-nsis-threats-03.txt
Expires: April 2004

October 2003

Security Threats for NSIS
<draft-ietf-nsis-threats-03.txt>

Status of this Memo

Abstract

This threats document provides a detailed analysis of the security threats relevant for the NSIS working group. It motivates and helps to understand various security considerations in the NSIS Requirements, Framework and Protocol proposals. This document does not describe vulnerabilities of specific NSIS protocols.

Table of Contents

1. Introduction

Whenever a new protocol has to be developed or existing protocols have to be modified their security threats should be evaluated. The process of securing protocols is separated into individual steps. To address security in the NSIS working group a number of steps have been taken:

- NSIS Analysis Activities (e.g. RSVP Security Properties)
- Security Threats for NSIS
- NSIS Requirements
- NSIS Framework
- NSIS Protocol Proposals

This document identifies the basic security threats that need to be addressed by the NSIS signaling protocol design. In addition, although the base protocol might be secure, some extensions may cause problems when used in a particular environment. Furthermore it is necessary to investigate the context in which a signaling protocol is used and the architecture where it is integrated. As an example of such interaction accounting and charging are taken into account in this document, since without an appropriate integration of the two it is difficult to deploy any NSIS solution. This interaction is also subject to discussion within the NSIS framework.

This document uses NSIS terms defined in [Bru03].

2. Relevant communication models

Signaling messages traverse different network parts, which demand different security protection and raise different security problems. The difference in security protection is mainly caused by the fact that the NSIS signaling messages cross trust boundaries where different trust relationships are prevalent. Often a categorization into first-peer/last-peer, intra-domain and inter-domain communication is applicable (see Figure 1).

The main properties of the listed network parts are briefly described in this section and the threats of Section 3 and Section 4 classify them to generic threats and signaling specific threats. Figure 1 depicts a typical end-to-end communication scenario including an access part between the NSIS end entities and the nearest NSIS hops, respectively. This "first-peer communication" commonly comes with specific security requirements (as described below), especially important for properly addressing security in mobile scenarios. Differences in the trust relationship and the required security for first-peer communication, compared to other parts of the signaling path, might exist.

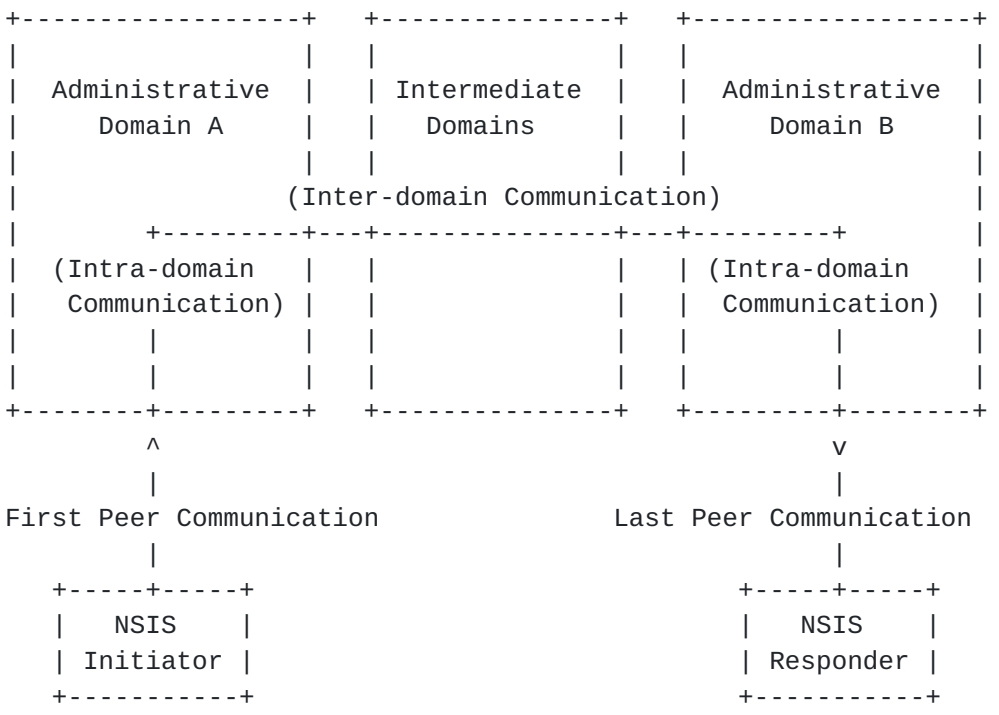
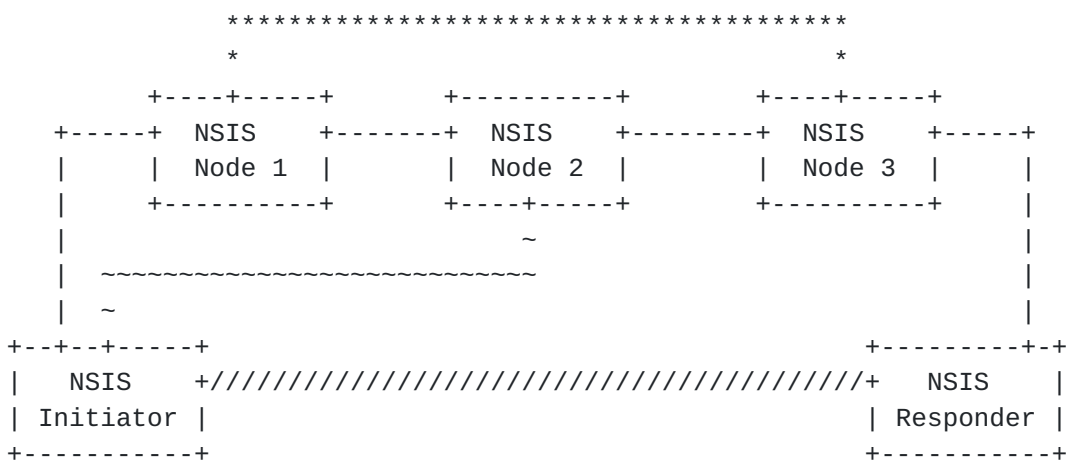


Figure 1: Involved Network Parts

To further refine the above differentiation based on network parts that NSIS signaling may traverse, we consider trust relationships between NSIS hops.
Additional threats may apply to NSIS communication where one entity involved is an end-entity (initiator or responder) and the other entity is any intermediate hop not being the first peer. This is typically called end-to-middle scenario. The motivation for including this configuration stems for example from the SIP [RFC3261] protocol.

To counter a number of specific security threats, any intermediate SIP hop may request a SIP end entity (UA) to authenticate. Such functionality in general seems to be useful for intermediaries at the borders of trust domains that signaling messages need to traverse.
Intermediate NSIS hops as well may have to deal with specific security threats that do not (directly) relate to end-entities. This scenario is called middle-to-middle. A typical example of middle-to-middle communication is between two NSIS hops at the border of their respective trust domains (i.e. inter-domain communication). NSIS messages may have to traverse one or more untrusted hops between these NSIS entities.
Figure 2 illustrates these additional scenarios. The first-peer case discussed further above is covered by the peer-to-peer trust relationships between end entity and closest hop, respectively.



Legend:
-----: Peer-to-Peer Trust Relationship
/////: End-to-End Trust Relationship
*****: Middle-to-Middle Trust Relationship
~~~~~: End-to-Middle Trust Relationship

Figure 2: Trust Relationships

2.1 First-Peer Communication

First peer communication refers to the peer-to-peer interaction between a signaling message originator, the NSIS Initiator (NI), and the first NSIS aware entity along the path. Assumptions about the threats, security requirements and the available trust relationships may be difficult here.
To illustrate this, in many mobility environments it is difficult to assume the existence of a pre-established security association directly available for NSIS peers involved in first-peer communication, as these peers cannot be assumed to have any relation between each other in advance. For enterprise networks, in contrast, the situation is different. Usually there is a fairly strong (pre-established) trust relationship between the peers. Enterprise network administrators usually have some degree of freedom to select the appropriate security protection and to enforce it. The choice of selecting a security mechanism is therefore often influenced by the already available infrastructure. Per-session negotiation of security mechanisms is therefore often not required (which, in contrast, is required for the mobility case).

For first-peer communication, especially threats related to initial security association setup, or threats due to replay attacks, lack of confidentiality, denial of service, integrity violation or identity spoofing are relevant, an potentially lead to theft of service and
~~~~~

fraud.

2.2 End-to-Middle Communication

End-to-middle interaction in signaling may be required to e.g. grant end-entities access to specific services in trust domains different from the one the first peer belongs to. Threats specific to this scenario may be introduced by untrusted intermediate NSIS hops that maliciously alter NSIS signaling. These threats are still relevant if security mechanisms are in place between the NSIS hops, but terminate at each hop (e.g. IPsec hop-by-hop protection).

2.3 Intra-Domain Communication

After having been verified at the first peer, an NSIS signaling message traverses the network within the same administrative domain the first peer belongs to. Since the request has already been authenticated and authorized threats are different to those described in the previous sections. To differentiate first-peer communication with the intra-domain communication (i.e. communication internally within one administrative domain) we assume that no end hosts have direct access to the internal network nodes, except the first peer. We furthermore assume that NSIS peers within the same administrative domain have at least some sort of trust relationship.

2.4 Inter-Domain Communication

The threat assumptions between the borders of different administrative domains largely depend on the authorization procedures. If one domain forges QoS reservations then this domain may also have to pay for the reservation. Hence in this case, there is no real benefit for this domain to forge a QoS reservation. If an end host is directly charged by intermediate domains (i.e. by a domain different from the malicious domain) such an attack may be quite a reasonable threat.

However, security protection of messages transmitted between different administrative domains is still necessary to tackle attacks like spoofing, integrity violation, or denial of service between these domains, e.g. to allow proper accounting. The number of neighboring domains is usually rather limited (compared to first-peer communication) which causes fewer problems for the key management required for securing inter-domain NSIS signaling.

Signaling information other than QoS service parameters such as policy rules in case of middlebox communication demands different assumptions for inter-domain communication. Trust assumptions and business relationships are of particular importance for their communication.

If signaling messages are conveyed transparently in the core network (i.e. they are not intercepted and processed in the core network) then the signaling message communication effectively takes place between access networks. This might place a burden on the key management infrastructure required between these access networks which might not know each other in advance. This might lead to an inability to secure signaling messages for a direct communication between the access networks. Hence, this can be seen as a serious deployment problem since it might be unacceptable for an access network service provider to perform processing (QoS reservations, policy rule installation at firewalls) triggered by unprotected incoming signaling messages.

2.5 End-to-End Communication

NSIS aims to signal information between the initiator and the responder. This section refers to the trust relationships required between the end points in cases where security protection is required. Note that this security protection is likely to be required only for certain objects such as pricing and charging related information. Protecting the entire signaling message is not possible since intermediate NSIS nodes need to (a) inspect various objects and (b) need to add, modify or delete objects from the signaling message.

The following example tries to illustrate a possible application of end-to-end protection for objects carried within the NSIS signaling protocol. Alice, the data sender, wants Bob, the data receiver, to pay for a QoS reservation (reverse charging). Bob wants to be assured that the QoS signaling message he receives was indeed transmitted by Alice because he is only willing to pay for particular users and not for everyone. Hence Bob wants to verify that the request came from Alice (authentication) and that the included parameters are unmodified. Additionally it might be necessary to secure a negotiation step and to secure deliver authorization information to the involved parties. Information which is required to compute an authorization decision (such as prices or QoS objects) also needs proper security protection.

Typical threats in such a scenario range from modification of QoS objects or price information (i.e. Bob has to pay more), fraud (i.e. to force Bob always to pay for the reservations) to identity spoofing (i.e. the adversary claims to be Alice).

Regarding end-to-end security one additional issue needs to be addressed - delegation. Whenever a signaling is addressed end-to-end and an arbitrary node along the path acts as a proxy on behalf of the other endpoint a delegation mechanism is required to provide secure interaction. This might lead to additional complexity.

2.6 Middle-to-middle Communication

We do not explicitly consider the middle-to-middle case here, although it is important, since it is already covered by either intra- or inter-domain communication depending on the location of the involved entities.

3. Generic Threats

This section provides threat scenarios that are applicable to signaling protocols. Note that some threat scenarios use the term user instead of NSIS Initiator. This is mainly because security protocols allow a differentiation between entities being hosts and users (based on the identities used).

3.1 Man-in-the-middle attacks

We differentiate this type of attack according to the separation of different steps, or phases, for securing protocols that is typically made. Therefore, this section starts with a brief motivation of this separation.

Security protection of protocols is often separated into two steps. The first step provides entity authentication and key establishment whereas the second step provides message protection using the previously established security association. The first step usually tends to be more expensive than the second which is also the main reason for separation. If messages are transmitted very infrequently then these two steps are collapsed into a single and usually rather costly step. One such example is e-mail protection via S/MIME. An example for a two-step approach is provided by IKE/IPsec. We use this separation to cover the different threats in more detail.
The first paragraph describes security threats where two peers do not already share a security association, or do not use security mechanisms at all. The next paragraph describes threats which are applicable when a security association is already established. Finally a denial of service attack is described which is applicable to a signaling message when no separation between SA establishment and signaling protection takes place. Particularly the discovery procedure is vulnerable against a number of attacks.

- Attacks during NSIS SA Establishment

During the process of establishing a security association an adversary fools the signaling message initiator with respect to the entity to which it has to authenticate. The man-in-the-middle adversary is able to modify signaling messages to mount DoS attacks. In addition, it may be able to terminate NSIS messages of the

Initiator and inject messages to a peer itself, therefore acting as the peer to the initiator and as the initiator to the peer. This results in the initiator wrongly believing that it talks to the "real" network whereas it is actually attached to an adversary. For this attack to be successful, pre-conditions have to hold which are described with the following two cases:

- Missing Authentication

The first case addresses missing authentication between the neighboring peers: Without authentication a NI, NR or NF is unable to detect an adversary. However, in some cases protection might be difficult to accomplish in a practical environment either because the next peer is unknown, because of misbelieved trust relationships in parts of the network or because of the inability to establish proper security protection (inter-domain signaling messages, dynamic establishment of a security association, etc.). If one of the communication endpoints is unknown then for some security mechanisms it is either not possible or very difficult to apply appropriate security protection. Sometimes network administrators use intra-domain signaling messages without proper security. Such a configuration would then allow an adversary on a compromised non-NSIS aware node to interfere with nodes running an NSIS signaling protocol. Note that this type of threat goes beyond a threat caused by malicious NSIS nodes (described in Section 4.7).

- Unilateral Authentication

In case of a unilateral authentication the NSIS entity that does not authenticate its peer is unable to discover the man-in-the-middle adversary. Although authentication of signaling messages should take place between each peer participating in the protocol operation special attention is given here to first-peer communication. Unilateral authentication between an end host and the first peer (just authenticating the end host) is still common today, but certainly opens up many possibilities for MITM attackers impersonating either the end host or the (administrative domain represented by the) first peer.

Missing or unilateral authentication, as described above, are a general problem of network access without appropriate authentication, and should not be considered as valid for the NSIS signaling protocol, only. Obviously there is a strong need to correctly address them in a future NSIS protocol. The signaling protocols addressed by NSIS are different to other protocols, where only two entities are involved. Note, that especially first-peer authentication is important, as the impacts of a security breach could impact nodes beyond the directly involved entities (or even beyond a local

network).

Finally it should be noted that the signaling protocol should be considered as a peer-to-peer protocol where the roles of initiator and responder can be reversed at any time. This leads to the conclusion that unilateral authentication is not very useful for such a protocol. However there might be a need to have some form of asymmetry in the authentication process whereby one entity uses a different authentication mechanism than the other one. As an example the combination of symmetric and asymmetric cryptography should be mentioned.

- Weak Authentication

This threat addresses weak authentication mechanisms whereby information transmitted during the NSIS SA establishment process may leak passwords and/or may allow offline dictionary attacks. This threat is applicable to NSIS for the process of selecting certain security mechanisms.

3.2 Adversary being able to replay signaling messages

This threat scenario covers the case where an adversary eavesdrops and collects signaling messages and replays them at a later point in time (or at a different place, or uses parts of them at a different place or in a different way - e.g. cut and paste attacks). Without proper replay protection an adversary might mount man-in-the-middle, denial of service and theft of service attacks.

A more difficult attack that may cause problems even in case of replay protection requires the adversary to crash an NSIS aware node to loose state information (sequence numbers, security associations, etc.) and to be able to replay old signaling messages. This attack addresses re-synchronization deficiencies.

3.3 Adversary being able to inject/modify messages

This type of threat addresses integrity violations whereby an adversary modifies signaling messages (e.g. by acting as a man-in-the-middle attacker) to cause an unexpected network behavior. Possible actions an adversary might consider for its attack are reordering, delaying, dropping, injecting and modifying.

An adversary may inject a signaling message requesting a large amount of resources (possibly using a different user identity). Other resource requests could then be rejected. In combination with identity spoofing it is also possible to accomplish fraud. This attack is only successful in absence of signaling message protection and authentication.

Some directly related threats are described in Section 4.7, 4.4 and 4.8.

3.4 Insecure Parameter Exchange/Negotiation

Protocols, which should be useful for a variety of scenarios, tend to have different security requirements. It is often difficult to meet these (sometimes conflicting requirements) with a single security mechanism or fixed security parameters. Often a selection of mechanisms and parameters are offered. Therefore a protocol exchange is required to agree on some security mechanisms/parameters. An insecure parameter exchange/negotiation protocol exchange can help an adversary to mount a downgrading attack by selecting weaker mechanisms than desired. Hence without protecting the negotiation process the security of an NSIS protocol might be as secure as the weakest mechanism if no configuration parameters (for example a security policy disallowing the weakest mechanism, etc.) are used otherwise.

4. Signaling specific Threats

This section lists both threats and attacks on the NSIS signaling protocol. A number of reasons might lead to an attack. Fraud is an example of an attack which might be caused by a number of reasons: missing replay protection, missing protection of authorization tokens, identity spoofing, missing authentication and many more might help an adversary to steal resources. These reasons which could lead to an attack are primarily addressed in this section.

In some cases we point to specific attacks which again might have a subsequent result since well-established security terms, such as denial of service, have to be addressed.

4.1 Threats based on NSIS SA Usage

Once a security association is established (and used to protect signaling messages) basic attacks are prevented. However, a malicious NSIS node is still able to perform various attacks as described in Section 4.7. Replay attacks can be a problem when an NSIS node crashes, restarts and performs state re-establishment. Proper re-synchronization capability of the security mechanism must therefore address this problem.

4.2 Threats based on combining Signaling and SA Establishment

These threats may lead to attacks which allow an adversary to flood an NSIS node with bogus signaling messages to cause a denial of service attack.

When a signaling message arrives at an NSIS aware network element some processing is required. If this message contains security objects such as digital signatures and no security association is already available then some processing is required for the cryptographic verification. Since NSIS signaling should not require several roundtrips between two NSIS peers it is difficult to provide DoS protection mechanisms commonly found in authentication and key agreement protocols. Signaling messages can be idempotent which means that they contain the same amount of information as the original message. An example would be a 'refresh' message which is in this case equivalent to a create message. This property enables that a refresh message creates new state along a new path although no previous state is available. In order for this to work it is necessary to use specific classes of cryptographic mechanisms supporting this behavior. An example is a digital signature based scheme which, however, should be used with care due to possible denial of service attacks. The problems of these types of message exchanges with public key based protection are described in [AN97] and in [ALN00].

Additionally to the threat described above an incoming signaling message might require time consuming processing (computations, state maintenance, timer setting, etc) and communication with third-party nodes including policy servers, LDAP servers, etc. If an adversary is able to transmit a large number of signaling messages (for example with QoS reservation requests) with invalid credentials then the verifying node may not be able to process further reservation messages by legitimate users.

Further threats could be introduced by allowing an adversary to gain additional information by injecting error messages or by forcing the creation of error messages.

4.3 Eavesdropping and Traffic Analysis

This section covers threats whereby an adversary is able to eavesdrop signaling messages. The collected signaling packets may serve for the purpose of traffic analysis or to later mount replay attacks as described in the Section 3.2. The eavesdropper might learn QoS parameters, communication patterns, policy rules for firewall traversal, policy information, application identifiers, user identities, NAT bindings, authorization objects and more.

The capability for an adversary to eavesdrop signaling messages might violate a users privacy preference particularly if authentication or authorization information (including policies and profile information) exchanged in an unprotected fashion.

Note, that the above threats are also applicable if the messages are integrity protected which is often considered sufficient for signaling protocols.

Since the NSIS protocol signals messages through a number of nodes it is possible to differentiate between nodes actively participating in the NSIS protocol and others who do not actively participate in the NSIS protocol. For certain objects or messages it might be desirable to permit actively participating intermediate NSIS nodes to eavesdrop. As a further extension it might be desired that only the intended end points (NSIS initiator and NSIS responder) are able to read certain objects.

4.4 Identity Spoofing

Identity spoofing relevant for NSIS, appears in two flavors: First, identity spoofing can appear during the establishment of a security association if based on a weak authentication mechanism.

Eve, acting as an adversary, claims to be the registered user Alice by spoofing the identity of Alice. Thereby Eve causes the network to charge Alice for the consumed network resources. This type of attack is possible if authentication is done based on a simple username identifier (i.e. in absence of cryptographic authentication) or if authentication is provided for hosts and multiple users have access to a single host. This attack could also be classified as theft of service.

An adversary is able to exploit the established flow identifiers (required for QoS and middlebox communication (Midcom) specific signaling protocols). Some identifiers such as IP addresses, transport protocol identifiers, port numbers, flow labels (see [RFC1809] and [RC+03]) and others are communicated in these protocols. Modification of these flow identifiers causes quality of service reservations or policy rules at middleboxes to be either ineffective or exploitable by adversaries. An adversary could mount an attack by modifying the flow identifier of a signaling message.

NSIS signaling messages contain some sort of flow identifier, which is associated with a specified behavior (e.g. a particular flow experiences QoS treatment or allows packets to traverse a firewall, etc.). An adversary might therefore use IP spoofing and inject data packets to benefit from previously installed flow identifiers.

The following threat is caused by identity spoofing of transmitted data traffic. The spoofed identity is thereby the source IP addresses. For this attack to be successful accounting records are collected based on the source IP address and not on a SPI due to

IPSec protection. After the network receives a properly protected

reservation request, transmitted by the legitimate user Alice, Traffic Selectors are installed at the corresponding devices (for example edge router). These Traffic Selectors are used for flow identification and allow to match data traffic originated from a given source address to be assigned to a particular QoS reservation. The adversary Eve now spoofs the IP address of the Alice. Additionally Alice's host may be crashed by the adversary as a result of a denial of service attack or lost connectivity for example because of mobility reasons. If both nodes are located at the same link and use the same IP address then obviously a duplicate IP address will be detected. Assuming that only Eve is present at the link then she is able to receive and transmit data (for example RTP data traffic), which receives preferential QoS treatment based on the previous reservation. Depending on the installed Traffic Selector granularity Eve might have more possibilities to exploit the QoS reservation or a pin-holed firewall. Assuming the soft state paradigm, where periodical refresh messages are required, the absence of Alice will not be detected until the next signaling message appears and forces Eve to respond with a protected signaling message. Again this issue is not only applicable to QoS traffic but the existence of QoS reservation causes more difficulties since this type of traffic is more expensive. The same procedure is also applicable to a Middlebox communication protocol.

The ability for an adversary to inject data traffic which matches a certain flow identifier established by a legitimate user often requires the ability to also receive the data traffic. This is, however, only true if the flow identifier consists of values which contain addresses used for routing. If we imagine to use attributes for a flow identifier where such a property is not required then identity spoofing and injecting traffic is much easier. An adversary can use a nearly arbitrary endpoint identifier to experience the desired result. Obviously the endpoint identifiers are still not irrelevant since the messages have to travel the same path through the network.

Data traffic marking based on DiffServ is such an example. Whenever an ingress router uses only marked incoming data traffic for admission control procedures then various attacks are possible. These problems are known in the DiffServ community for a long time and documented in various DiffServ related documents. The IPSec protection of DiffServ Code Points is described in Section 6.2 of [RFC2745]. Related security issues (for example denial of service attacks) are described in Section 6.1 of the same document.

4.5 Missing Protection of Authorization Information

Authorization is an important step for providing resources such as

QoS reservations, NAT bindings and pinholes on firewalls.

Authorization information might be delivered to the NSIS participating entities in a number of ways.

Typically the authenticated identity is used to assist during the authorization procedure as e.g. described in [RFC3812]. Depending on the chosen authentication protocol certain threats may exist. Section 3 discusses a number of issues related to this approach when the authentication and key exchange protocol is used to establish session keys for signaling message protection.

Another approach is to use some sort of authorization token. The functionality and structure of such an authorization token for RSVP is described in [RFC3520] and in [RFC3521].

The interaction between different protocols based on authorization tokens, however, requires some care. By using such an authorization token it is possible to link state information between different protocols. Returning an unprotected authorization token to the end host might allow an adversary (for example an eavesdropper) to steal resources. An adversary might also use the token to learn communication patterns. An untrustworthy end host might also modify the token content.

The Session/Reservation Ownership problem can also be considered as an authorization problem. Details are described in Section 4.10. In enterprise networks authorization is often coupled with membership to a particular class of users/groups. This type of information can either be delivered as part of the authentication and key agreement procedure or has to be retrieved via separate protocols from other entities. If an adversary manages to modify information relevant for determining authorization or the outcome of the authorization process itself then theft of service might be the consequence.

4.6 Missing Non-Repudiation

Repudiation in this context refers to a problem where one party later denies to have requested a certain action (such as a QoS reservation). The problem of a missing non-repudiation property appears in two flavors:

From a service provider point-of-view the following threat may be worth an investigation. A user may deny to have issued reservation request for which it was charged. A service provider may then like to prove that a particular user issued reservation requests.

The same threat can be interpreted from the user's point-of-view. A service provider claims to have received a number of reservation requests. The user in question thinks that he never issued those

requests and wants to have a proof for correct service usage for a given set of QoS parameters.

In today's telecommunication networks non-repudiation is not provided. The user has to trust the network operator to correctly meter the traffic, collect and merge accounting data and that no unforeseen problems occur. If a signaling protocol is used to establish QoS reservations with the non-repudiation property for the authorized resources then it has an impact on the protocol design.

Non-repudiation poses additional requirements on the security mechanisms as it can only be provided through public-key cryptography. As this would often increase the overall cost for security, threats related to missing non-repudiation are only considered relevant for certain specific scenarios (e.g. specific authorization mechanisms) and not for general NSIS signaling.

4.7 Malicious NSIS Entity

Network elements within a domain (intra-domain) experience a different trust relationship with regard to the security protection of signaling messages compared to the edge NSIS entity. We assume that edge NSIS entity have the responsibility to perform cryptographic processing (authentication, integrity and replay protection, authorization and accounting) for signaling message arriving from the outside. This prevents signaling messages to appear unprotected within the internal network. If, however, an adversary manages to take over an edge router then the security of the entire network is affected. An adversary is then able to launch a number of attacks including denial of service, integrity violation, replay, reordering and deletion of data packets and various other attacks. In case of policy rule installation a rogue firewall can cause harm to other firewalls by modifying the policy rules accordingly. The chain-of-trust principle applied in the peer-to-peer security protection cannot provide protection against a malicious NSIS node. An adversary with access to an NSIS router is then also able to get access to security associations to transmit secured signaling messages. Note that even non peer-to-peer security protection might not be able to fully prevent this problem. Since an NSIS node might issue signaling messages on behalf of someone else (by acting as a proxy) additional problems are the consequence.

An NSIS aware edge router is a critical component that requires strong security protection. A strong security policy applied at edge does not imply that all routers within an intra-domain network do not need to cryptographically verify signaling messages. If the chain-of-trust principle is deployed then the security protection of the entire path (in this case within the network of a single

administrative domain) is as strong as the weakest link. In our case

the edge router is the most critical component of this network that may also act as a security gateway/firewall for incoming/outgoing traffic. For outgoing traffic this device has to act according to the security policy of the local domain to apply the appropriate security protection.

For an adversary to mount this attack either an existing NSIS aware node along the path has to be successfully attacked or an adversary succeeds to convince another NSIS node to be the next NSIS peer (man-in-the-middle attack).

4.8 Denial of Service Attacks

A number of denial of service attacks can cause NSIS nodes to malfunction. Other attacks that could lead to DoS, such as man-in-the-middle attacks, replay attacks, injection or modification of signaling messages etc., are mentioned throughout this document.

- Path Finding

This threat tries to address potential denial of service attacks when the reservation setup is split into two phases i.e. path and reservation (as for example used in receiver based reservation setup). For this example we assume that the node transmitting the path message is not charged for the path message itself and is able to issue a high number of reservation requests (possibly in a distributed fashion). Charging is activated only after successful verification of the reservation request. The reservations are however never intended to be successful because of various reasons: the destination node cannot be reached; it is not responding or simply rejects the reservation. An adversary can benefit from the fact that state has already been allocated along the path for various processing tasks including path pinning.

- Discovery Phase

Signaling information to a large number of entities along a data path requires some sort of discovery. This discovery process is vulnerable to a number of attacks since it is difficult to secure. An adversary can use the discovery mechanisms to convince an entity to signal information to another entity which is not along the data path or to cause the discovery process to fail. In the first case the signaling protocol could be correctly continued with the problem that policy rules are installed at incorrect firewalls or QoS resource reservations take place at the wrong entities. For an end host this means that the protocol failed for unknown reasons.

- Faked Error/Response messages

An adversary may be able to inject false error/response messages as part of a denial of service attack. This could be either at the message signaling protocol level (NTLP), at the level of each client layer protocol (NSLP: QoS, Midcom, etc.) or at the transport level protocol. An adversary might cause unexpected protocol behavior, or might succeed with denial of service attacks. Especially the discovery protocol shows vulnerabilities with regard to this threat (see above discussion on discovery). In case that no separate discovery protocol is used by addressing signaling messages to end hosts only (with a Router Alert Option to intercept message as NSIS aware nodes) then an error message might be used to indicate a path change. Such a design is a combination of a discovery protocol together with a signaling message exchange protocol.

4.9 Disclosing the network topology

In some architectures there is a desire not to reveal the internal network structure (or other related information) to the outside world. An adversary might be able to use NSIS messages for network mapping (e.g. discovering which nodes exist, which use NSIS, what version, what resources are allocated, capabilities of nodes along a paths etc.). Discovery messages, traceroute, diagnostic messages (see [RFC2745] for a description of diagnostic message functionality for RSVP), query messages in addition to record route and route objects provide the potential to assist an adversary. Hence the requirement of not disclosing a network topology might conflict with another requirement to provide means for automatically discovering NSIS aware nodes or to provide diagnostic facilities (used for network monitoring and administration).

4.10 Missing protection of Session/Reservation Ownership

Figure 3 shows an NSIS Initiator which established state information at NSIS nodes along the path as part of the signaling procedure. As a result the Access Router1 Router 3 and Router 4 (and other nodes) store session state information including the Session Identifier SID-x.

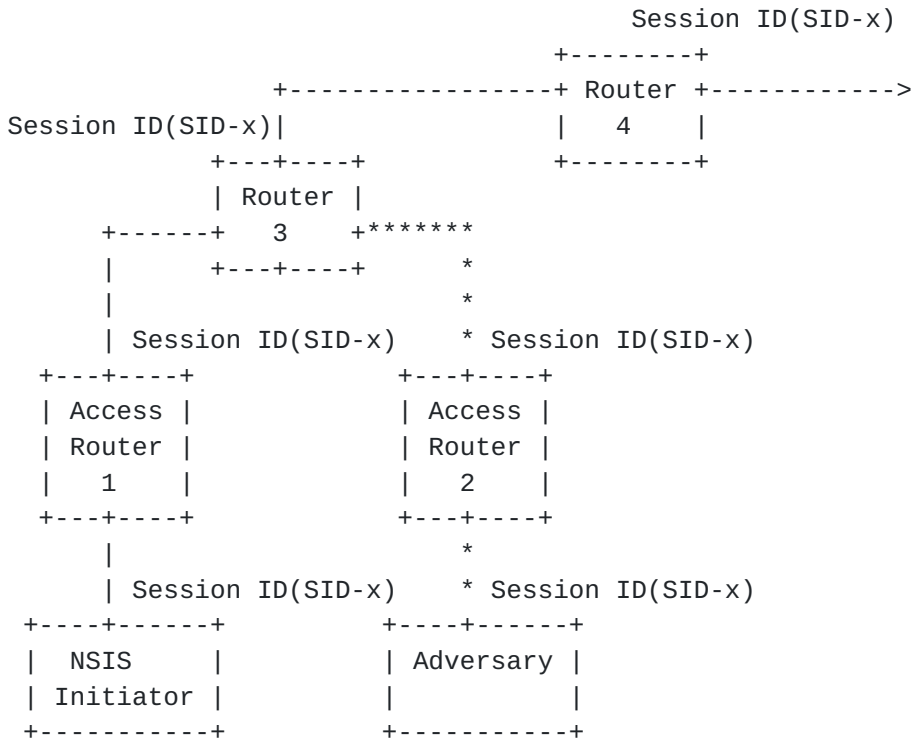


Figure 3: Session/Reservation Ownership

The Session Identifier is included in signaling messages to reference to the established state.

If an adversary was able to obtain the Session Identifier for example by eavesdropping signaling messages it is able to add the same Session Identifier SID-x to a new signaling message. When the signaling message hits Router3 (as shown in Figure 3) then existing state information can be modified. The adversary can then modify or delete the established reservation causing unexpected behavior for the legitimate user.

The source of the problem is that Router3 (cross-over router) is unable to decide whether the new signaling message was initiated from the owner of the session/reservation.

In addition, not only the initial signaling message originator is allowed to signal information during the lifetime of an established session. As part of the protocol any NSIS aware node along the path (and the path might change over time) could initiate a signaling message exchange. It might, for example, be necessary to provide mobility support or to trigger a local repair procedure. If only the initial signaling message originator is allowed to trigger signaling message exchanges some protocol behavior would not be possible.

In case that this threat is not addressed an adversary can launch denial of service, theft of service, and various other attacks.

4.11 Attacks against the transport mechanism

In [BL01] a two-level architecture is proposed which suggests to split an NSIS protocol into layers: a signaling message transport specific layer and an application specific layer. This architectural assumption is also considered within the NSIS framework [HF+03]. Most of the threats described in this document are applicable to the application specific part for signaling QoS or middlebox specific information. There are, however, some threats which are applicable to the transport of signaling messages.

Network or transport layer protocols lacking protection mechanisms are vulnerable to certain attacks such as header manipulation, DoS, spoofing of identities, session hijacking, unexpected aborts etc.

Malicious nodes can attack the congestion control mechanism to force NSIS nodes into a congestion avoidance state.

In case that an existing protocol is used for exchanging NSIS signaling messages then threats known from these protocols are relevant.

5. Security Considerations

This entire memo discusses security issues relevant for NSIS. To counter these threats security requirements have been listed in [Brun03]. Framework relevant topics have been incorporated into [HF+03].

6. Normative References

[Brun03] M. Brunner, "Requirements for QoS signaling protocols," Internet Draft, Internet Engineering Task Force, August 2003. Work in progress.

7. Informative References

[HF+03] R. Hancock, I. Freytsis, G. Karagiannis, J. Loughney, and S. V. den Bosch, "Next steps in signaling: Framework," Internet Draft, Internet Engineering Task Force, September 2003. Work in progress.

[RFC1809] C. Partridge, "Using the flow label field in IPv6," RFC 1809, Internet Engineering Task Force, June 1995.

[RFC2745] A. Terzis, B. Braden, S. Vincent, and L. Zhang, "RSVP Diagnostic Messages," RFC 2745, Internet Engineering Task Force, Jan. 2000.

[RFC3182] Yadav, S., Yavatkar, R., Pabbati, R., Ford, P., Moore, T., Herzog, S., Hess, R.: "Identity Representation for RSVP", RFC 3182, October, 2001.

[RFC3261] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: session initiation protocol," RFC 3261, Internet Engineering Task Force, June 2002.

[RFC3521] L. Hamer, B. Gage, and H. Shieh, "Framework for session set-up with media authorization," RFC 3521, Internet Engineering Task Force, April 2003.

[RFC3520] L. Hamer, B. Gage, B. Kosinski, and H. Shieh, "Session Authorization Policy Element", RFC 3520, Internet Engineering Task Force, April 2003.

[RC+03] J. Rajahalme, A. Conta, B. Carpenter, and S. Deering, "IPv6 Flow Label Specification," Internet Draft, Internet Engineering Task Force, April 2003. Work in progress.

[BL01] B. Braden and B. Lindell, "A two-level architecture for internet signaling," Internet Draft, Internet Engineering Task Force, Nov. 2001. (Expired).

[AN97] T. Aura and P. Nikander: "Stateless Connections", In Proceedings of the International Conference on Information and Communications Security (ICICSÆ97), Lecture Notes in Computer Science 1334, Springer, 1997.

[ALN00] T. Aura, J. Leiwo and P. Nikander: "Towards Network Denial of Service Resistant Protocols", In Proceedings of the 15th International Information Security Conference (IFIP/SEC 2000), Beijing, China, August 2000.

Acknowledgments

We would like to thank (in alphabetical order) Marcus Brunner, Jorge Cuellar, Mehmet Ersue, Xiaoming Fu and Robert Hancock for their comments to an initial version of this draft. Jorge and Robert gave us an extensive list of comments and provided information on additional threats.

Jukka Manner, Martin Buechli, Roland Bless, Marcus Brunner, Michael Thomas, Cedric Aoun, John Loughney, Rene Solwitsch, Cornelia Kappler, and Mohan Parthasarathy provided comments to a recent version of this draft. Their input helped to improve the content of this document. Particularly Roland Bless, Michael Thomas and Cornelia Kappler provided good proposals for regrouping and restructuring.

Author's Addresses

Hannes Tschofenig
Siemens AG
Corporate Technology
CT IC 3
Otto-Hahn-Ring 6
81739 Munich
Germany
EMail: Hannes.Tschofenig@siemens.com

Dirk Kroeselberg
Siemens AG
Otto-Hahn-Ring 6
81739 Munich
Germany
EMail: Dirk.Kroeselberg@siemens.com

Full Copyright Statement

Acknowledgement

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