3GPP TSG SA WG3 Security — SA3#35 **S3-040720**
October 5-8, 2004
St Paul's Bay, Malta

Source: **BT Group**
Contacts: **Colin Blanchard colin.blanchard@bt.com,**
Martin Moore martin.moore@bt.com
Title: **Proposal for an informative Annex to the 3GPP TS 33.203 on support of end user devices behind a NA(P)T firewall and protection of RTP media flows**

Document for: **Discussion and decision**
Agenda Item: **TBA**

1. Introduction

When the decision was taken by SA3 to base IMS security on IPSec, it was acknowledged that the solution would not support a UE behind a NAT firewall router. This is not an issue with the present IMS architecture in 3GPP, but may impose restrictions, as the 3GPP network and service develops. It also presents difficulties with reusing the 3GPP IMS security specification in other contexts, such as WLAN and ADSL access networks, which are likely to incorporate NAT between the end-user device and the SIP server. However, this paper shows that the existing standard (TS33.203) can address 1) future 3GPP architectures and 2) Broadband access networks, without the need for a SIP application layer proxy (ALG) on the home gateway. In the Broadband case, the ability to benefit from a proven access security mechanism, without the need for changes to broadband routers already deployed, and with only minimal additions to the existing standards, will be immensely useful to network operators.

SA3 also took the decision to concentrate on providing protection for SIP signalling, and not the associated media streams, relying on the confidentiality protection provided by the underlying GPRS bearer instead. The paper also shows that although IMS access security standard TS33.203 does not currently support protection of the media streams, it may be extended to support separate IPSec access tunnels for SIP signalling and RTP media flows, as required by future architectures that have separate SIP and RTP proxies on the provider network.

It is intended that this paper forms the basis for an Informative Annex to the 3GPP TS 33.203 Access Network Security standard, as no normative changes to the specification are considered necessary at this stage. The future adoption of a SIP application layer proxy, for example when the access gateway employs QoS mechanisms, is in no way precluded by this contribution, although considerable further work would be needed to support a SIP application layer proxy as an adjunct, and would have an impact on the normative part of TS33.203.

2. Background

The SIP protocol, on which IMS is based, is particularly vulnerable. SIP allows an end user to signal directly into the core network, which no longer benefits from the insulation of a separate access signalling network. SIP has little by way of inherent protection. 3GPP security standards provide for the strong authentication and encryption of SIP traffic across the access network, between SIP client and network proxy, and between core network proxies. This is often referred to as “hop-by-hop IPsec”. The end user can be sure that their sessions are adequately protected (opportunities for interception, insertion and replay attacks are minimised). The network operator can be sure that attacks on the signalling network, and on the services that use it, are far harder to mount. The network proxy (P-CSCF in 3GPP parlance) can also have a hiding role, if it terminates a user SIP session, and associated RTP media streams, and initiates a new session with the next-hop proxy. The called party will be unaware of the true identity of the caller.

3GPP has chosen to use IPsec for securing both IMS access and network domain connections, in hop-by-hop fashion, rather than Transport Layer Security (TLS). TLS is an IETF standard requiring the use of Public Key Certificates for authentication (not always feasible for access networks). TLS cannot be used to secure UDP traffic, notably RTP media streams.

3GPP standards were designed with 3rd Generation Mobile (UMTS) networks in mind. However, the use of IMS now extends well beyond UMTS. For example, in the fixed network, the access architecture will be based on Multi-Service Access Nodes (MSANs), which are designed to support a wide variety of access media and mechanisms. In particular, residential and business customers with xDSL broadband lines must be catered for. Typically, a broadband customer will have a Network (or Port) Address Translating (NAT) firewall router, with built-in DHCP support for the addressing of hosts behind the NAT. The 3GPP access network security standard does *not* currently support devices behind a NAT firewall. This paper shows how, with only minimal additions, the existing standard can be extended to cover broadband access networks.

The existence of NAT within the core network, or multiple instances of NAT in the access network, is outside the scope of this discussion paper.

3. Solution

3.1 Network Setup and Components

A working demonstrator has been built, to prove the concepts outlined here. Only one SIP client is behind a NAT, in order to demonstrate operation with and without NAT, but there is no reason why the NAT traversal mechanism could not operate on both access links.

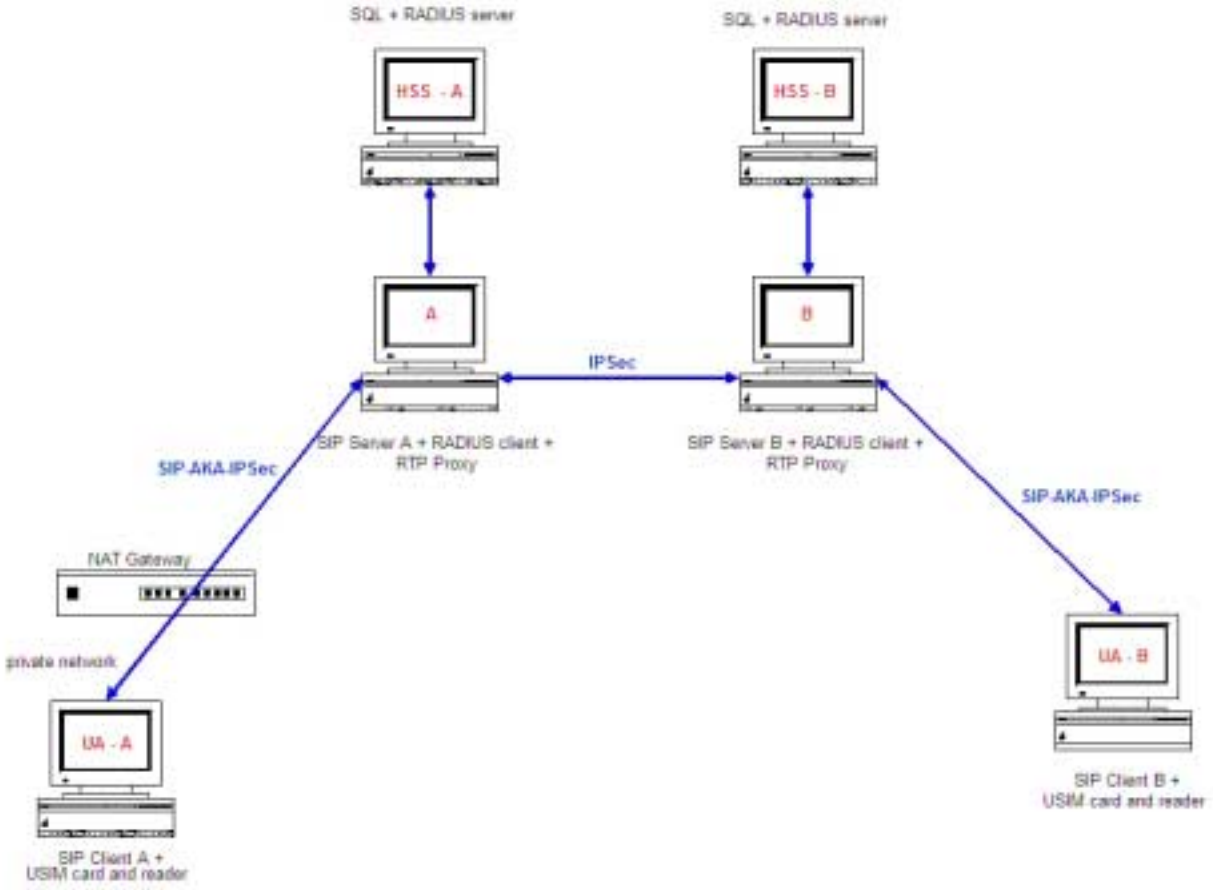

Figure 1 Logical representation of the secure SIP network set-up

For the purposes of this discussion, it is easier to think of the set-up in terms of IPsec endpoints, as in **Figure 1**
In reality, in the demonstrator, the interfaces on all the machines have routable IP addresses on a subnet of the local LAN, with the exception of the machine behind the NAT gateway, which has an RFC 1918 private address. The NAT gateway could be any standard commercial offering, which supports DHCP private addressing, NAT and routing.

To support end user devices behind a NAT firewall and protection of RTP media flows, however, the software components on each node have to be implemented in a specific way. It must be stressed that this does *not* imply normative changes to the underlying standard. These implementation considerations are now described for each node in the logical representation of **Figure 1**.

- **SIP Server A:** SIP Server in Network A, The following components need to be installed on this node:

    1. SIP Proxy
    2. Software support for SIP-AKA authentication[1]
    3. Radius client, which has, support for AKA and can communicate with the HSS
    4. An IPSec client with the necessary hooks to build an IPSec security association, once the SIP client has been successfully authenticated.
       *Note that it may be necessary to extend the proxy to allow the building and updating of an IPSec SA. Many do not have built-in support for IPsec.
       **Also note that it may be necessary to extend the IPsec client, to support UDP encapsulation for clients behind NAT[2]. UDP encapsulation is normally linked to Internet Key Exchange[3] (IKE, the usual mechanism for key management in IPsec). IKE is not used in 3GPP IMS security, so UDP encapsulation must be implemented within AKA instead.
       ***TS33.203 mandates the use of transport mode IPsec. While this will work with UDP encapsulation for NAT traversal, it is recommended that tunnel mode IPsec, with DHCP or similar mechanism for assignment of inner IP address, be used in broadband access environments. This is to avoid conflicts, when NAT routers assign addresses from the same range, e.g. 192.168.0.0
       ****While using tunnel mode IPSec and UDP encapsulation for clients behind NAT, the SIP server needs to see the internal address of the tunnel. Hence once the IPSec SA is established, any NAT traversal support on the SIP server needs to be turned off. Only the initial *unencrypted* registration (SIP-AKA) messages require the NAT traversal support. The SIP server may need to be extended to support this requirement.
    5. Support for proxying the RTP media traffic.
    6. Support for long-standing IPSec security associations with other SIP related network entities like the HSS and other SIP servers. How these network, as opposed to access level, IPSec SAs are built is assumed to be covered by 3GPP Network Domain security, and are not considered further in this paper.

- **SIP Server B:** This machine is an exact replica of SIP Server A, the only difference being that it represents a different operator or SIP domain

- **HSS A and HSS B:** These servers store the subscriber information, including the master secret that the RADIUS server will use to generate the AKA quintuplets. No changes are required to support the additional features of NAT traversal and access link protection. However, the Generic Authentication Architecture (GAA) concept[4], and the BootStrapping Function (BSF) within it, could be used to provide *separate* secure IPsec access tunnels for SIP messaging and RTP media flows, for each host requiring SIP services, This would require the implementation of the Zn interface to the BSF, rather than a direct connection to the HSS.

- **SIP Client A:** This machine sits behind a NAT gateway, and requires a SIP Client with support for SIP-AKA authentication and an IPSec client with support for UDP encapsulation, and with the necessary hooks to build an IPSec SA once the SIP Client has been successfully authenticated. It may be necessary to extend the SIP client to allow the building and updating an IPSec SA, if IPsec client functionality is not built-in. The following protocol will also need to be implemented:

    1. When the access link IPSec SAs have been successfully set up at both ends, SIP Client A shall send at least one packet (ICMP ping for example) to the SIP server. This will create a pinhole on the NAT gateway, which will allow UDP encapsulated IPSec traffic, coming from the SIP Server, to pass

---

[1] TS33.102 3GPP Security Architecture, Section 6.3 Authentication and Key Agreement
TS35.205-208 Milenage example algorithm set

[2] Draft-ietf-ipsec-ipsec-udp-encaps-09.txt "UDP Encapsulation of IPsec ESP Packets"

[3] Draft-ietf-nat-t-ike-08.txt "Negotiation of NAT-Traversal in the IKE"

[4] TS33.220 3GPP Generic Authentication Architecture

through the NAT gateway. In the absence of a NAT pinhole, any incoming packets would be dropped by the gateway (unknown destination).

2. Once the pinhole is created, it may be kept alive by periodic keep-alive messages (either a SIP layer message, or an ICMP packet) sent either by SIP Server A, or by SIP Client A.
3. Note that since IPSec UDP encapsulation is used for NAT traversal, once the IPSec SA is established, any SIP/RTP layer support for NAT traversal is no longer required. Consequently, symmetric RTP support, for example, is no longer required (use of the same port number for sending and receiving the RTP traffic). Symmetric RTP is a proposed NAT traversal solution for media traffic, in the absence of any NAT traversal support from the lower layers.

- **SIP Client B:** This machine is an exact replica of SIP Client A, but is identified by a different SIP URL, and has a routable IP address. Thus NAT traversal support is unnecessary. SIP Server B is the registration server for client B.

No changes were required to the UICC and USIM application, or to the IMS application itself. (A video-calling application was used for the purposes of this demonstration, but any SIP-based application would suffice.)

3.2 Operation without NAT

Figure 2 describes how SIP-AKA works between SIP Client B and SIP Server B, when there is no NAT gateway between the two.

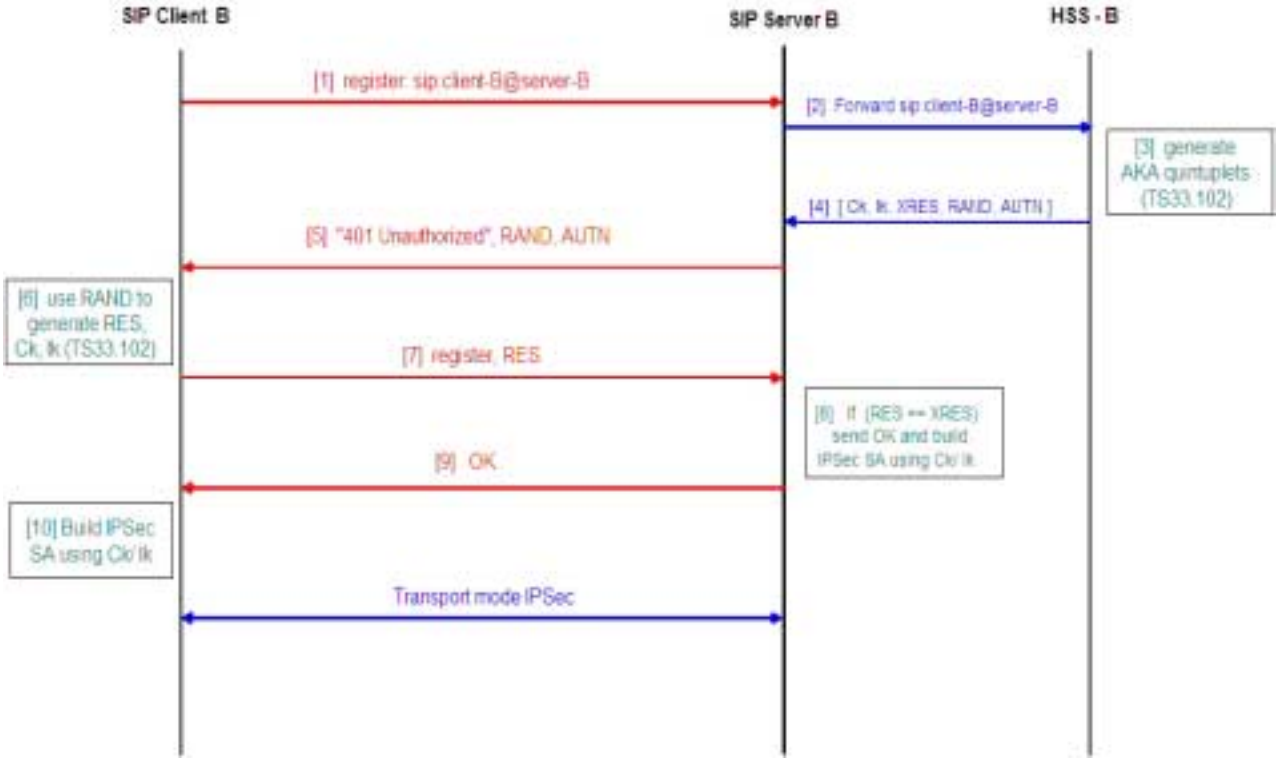

Figure 2 SIP-AKA without NAT

Step-by-step explanation:

1. SIP Client B sends a SIP REGISTER message to SIP Server B.
2. SIP Client B hasn't been authenticated. Therefore SIP Server B forwards the SIP Client B SIP URL to HSS B, via radiusclient[6].
3. HSS B uses the SIP URL to locate the user in its database, and to generate the necessary AKA quintuplets.
4. HSS B returns {CK, IK, XRES, RAND, AUTN} quintuplet to SIP Server B.

5. SIP Server B sends a “401 Unauthorized” SIP message to SIP Client B, along with RAND and AUTN which it received from HSS B.
6. The UMTS SIM card on UA B uses RAND to generate CK, IK and RES.
7. SIP Client B resends the SIP REGISTER message, along with RES to SIP Server B.
8. SIP Server B compares if RES is equal to XRES.
9. If they are equal, SIP Server B sends an OK message to SIP Client B and builds the required IPSec SA, using CK/IK keying material.
10. SIP Client B, on receiving the OK message, builds the corresponding IPSec SA in the reverse direction, using the same CK/IK keying material.

When the registration period expires, the user agent *must* re-register. So the same process as shown in Figure 2 is repeated, but this time all the SIP registration traffic goes encrypted. A successful re-registration results in both ends updating the IPSec Security Associations. When the SIP Client de-registers, IPSec Security Associations are deleted at both the client and the server ends.

3.3 Operation with NAT

Figure 3 describes how SIP-AKA works between SIP Client A and SIP Server A, when there is a NAT Gateway between the two.

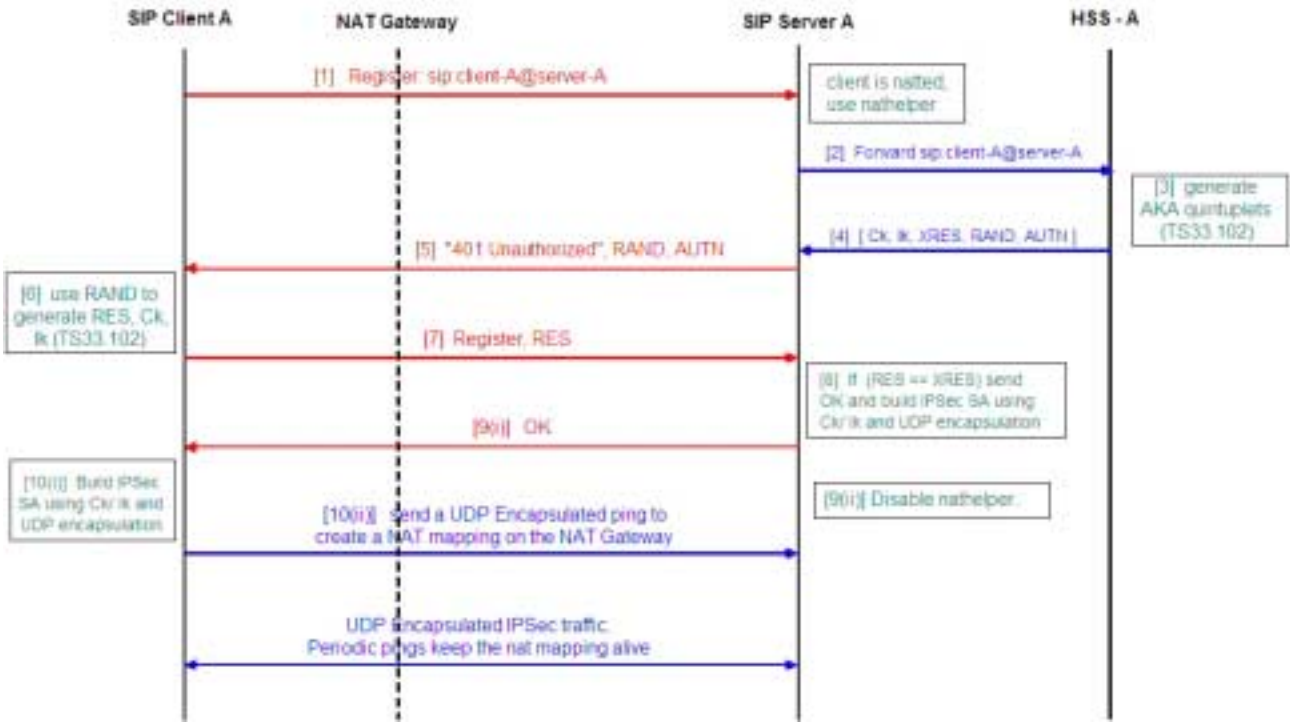

Figure 3 SIP-AKA with NAT

Step-by-step explanation:

1. SIP Client A sends a SIP REGISTER message to SIP Server A
2. SIP Server A sees a request coming in from a client behind a NAT (IP address in header and SIP contents do not coincide), and so turns on the nathelper[5] functionality for this client. As the client has not been authenticated, SIP Server A again forwards the SIP Client A SIP URL to HSS A via radiusclient[6].

[5] Our SIP servers were built on Linux. Nathelper is a Linux application. Any equivalent application may be used.

[6] radiusclient is a Linux application. Any equivalent RADIUS application may be used.

3. HSS A uses the SIP URL to locate the user in its database, and to generate the necessary AKA quintuplets.
4. HSS A returns {CK, IK, XRES, RAND, AUTN} to SIP Server A.
5. SIP Server A sends a “401 Unauthorized” SIP message to UA A, along with RAND and AUTN which it received from HSS A.
6. The UMTS SIM card on SIP Client A uses RAND to generate CK, IK and RES.
7. SIP Client A resends the SIP REGISTER message along with RES to SIP Server A.
8. SIP Server A compares RES to XRES (the expected response).
9. (i) If they are equal, SIP Server A sends an OK message to SIP Client A, and builds the required IPSec SA, using CK/IK keying material. An important distinction is that UDP encapsulation for NAT traversal is now turned on.
   (ii) SIP Server A switches off nathelper for SIP Client A.
10. (i) SIP Client A on receiving the OK message, builds the IPSec SA, using the same CK/IK keying material.
    (ii) SIP Client A sends a UDP-encapsulated ICMP ping to SIP Server A. This will create a pinhole in the NAT gateway, which will allow any incoming packets to get through the NAT. A keep-alive message sent periodically by SIP Client A keeps the NAT mapping alive. Note that the keep-alive message can come from either end. In this implementation the client is the sender.

When the registration period expires, the SIP Client must re-register. So the same process as shown in **Figure 3** is repeated, but this time all the SIP registration traffic is encrypted. A successful re-registration results in the both ends updating their respective IPSec security associations. When the SIP Client de-registers, the IPSec security association is deleted at both the client and the server ends.

4. The SIP application layer proxy as an additional option

The motivation behind this contribution was that in the broadband case, the ability to benefit from a proven access security mechanism, without the need for changes to broadband routers already deployed, and with only minimal additions to the existing standards, would be immensely useful to network operators. However, it has been suggested that management of QoS will require a SIP application layer proxy and if, as intended, the QoS mechanisms in question are application-driven, substantial changes to the TS33.203 will be required. A number of issues must be addressed:

- The end-to-end, access gateway-to-VoIP device, security model will be broken. This is a fundamental tenet of GSM and UMTS networks. An Application Layer Gateway (ALG) can only function if the payloads of IP packets are neither authenticated nor encrypted, so the end user must be prepared to delegate security to an intervening device, making access network connections on their behalf. This may be an issue for customers of NGN networks roaming to other home networks.
- In order to maintain IPsec-level access network protection, the home gateway must itself be IPsec-capable. Configuring and managing separate IPsec access tunnels on behalf of multiple users on the home network will not be easy.
- The ALG must be closely integrated with the firewall and NAT, if unacceptable security vulnerabilities are to be avoided.
- It must not be possible to mount Denial of Service (DoS) attacks on the NGN network, by manipulating the built-in QoS tools.
- A mechanism for securely managing what will inevitably be a more sophisticated device in the home must exist.

Future contributions will address this additional option.