3GPP TR 33.802 V0.2.0 (2005-11)

Technical Report

3rd Generation Partnership Project; Technical Specification Group Services and System Aspects Feasibility study on IMS Security Extensions (Release 7)

Keywords
IMS, security, IP, Multimedia, SIP

3GPP

Postal address

3GPP support office address
650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet
http://www.3gpp.org

Contents

Foreword 6
Introduction 6
1 Scope 7
2 References 7
3 Definitions, symbols and abbreviations 8
3.1 Definitions 8
3.2 Symbols 8
3.3 Abbreviations 8
4 Requirements 8
4.1 ETSI TISPAN NGN R1 Requirements 8
4.2 ETSI TISPAN NGN R2 Requirements 8
5 Analysis 8
5.1 IMS access security solution for NAT/FW traversal 8
5.1.1 Introduction 8
5.1.2 Alternatives 9
5.1.2.1 TLS 9
5.1.2.2 IPSec tunnel mode (terminating to P-CSCF) with UDP encapsulation 9
5.1.2.3 WLAN-IW scenario 3 with IMS Access Security (TS33.203) 10
5.2 NAT device traversal and interoperability issues for IMS Rel-7 10
5.2.1 Introduction 10
5.2.2 Discussion 10
5.2.2.1 NAT-related difficulties 10
5.2.2.2 Solutions in the protocols 11
5.2.2.3 Solutions in the NAT device 12
5.2.2.4 Multiple clients behind one NAT device 12
5.3 Analysis of Enhanced 3GPP R5/6 Access Security Mechanism 12
5.3.1 Introduction 12
5.3.2 Analysis 13
5.3.2.1 Implementation 13
5.3.2.2 Co-existence of Applications 15
5.3.2.3 Immature standards 15
5.3.2.4 Poor compatibility with NATs 15
5.4 Comparison of solutions for IMS signalling protection that are proposed to be added in 3GPP Release 7 (SA3 drafting Group at SA3 #41) 16
5.4.1 Introduction 16
5.4.2 Comparison 16
5.5 Signalling Protection 18
5.6 Media Protection 18
6 Conclusions 18
Annex A: Approaches to TLS based IMS security solutions 20
A.1 Introduction 20
A.2 Problem statement 20
A.3 Solution 20
A.3.1 IMS AKA for authentication 20
A.3.2 TLS protection 20
A.3.3 IMS AKA asserted TLS (I-TLS) signalling protection solution 21
A.3.3.1 Overview of IMS AKA asserted TLS solution 21
A.3.3.2 Interoperability with IMS Rel-5/6 23
A.3.3.2.1 Using security agreement 23
A.3.3.2.2 Fallback to Rel-5/6 24

A.3.3.3 The details of the token 25
A.3.4 PSK TLS for signalling protection 26
A.4 Discussion 28
A.5 References 29
Annex B: Enabling NAT traversal for signalling messages in the IMS access security framework 29
B.1 Introduction 29
B.2 Overview 30
B.2.1 Requirements and Objectives 30
B.2.2 Assumptions 30
B.2.3 Solution Outline 30
B.3 Detailed Solution Description 31
B.3.1 General problems with SIP and NAT (not specific to security) 31
B.3.2 NAT traversal for unprotected messages (not security-specific) 32
B.3.3 Detection of NAT traversal capabilities and presence of a NAT (partly security-specific) 32
B.3.4 NAT traversal of protected messages (security-specific) 33
B.3.4.1. UDP encapsulation using transport mode 33
B.3.4.2 UDP encapsulation using tunnel mode 35
B.3.5 Registering a Contact and routing of UE terminating requests (partly security-specific) 36
B.3.6 Keeping the NAT binding alive (not security-specific) 36
B.4 Establishing IPSec SAs and handling of UDP encapsulation 37
B.4.1 Using a separate UDP encapsulation function and UDP encapsulated tunnel mode 37
B.4.2 Using IPSec built-in UDP encapsulation features and UDP encapsulated transport mode 39
B.5 Multiple UEs behind the same NAT 40
B.5.1 Implications from the use of a common (public) IP address for multiple UEs 40
B.6 References 40
Annex C: Generic Network Tunnel (GNT) for NGN 41
Annex D: Enabling NAT traversal for signaling messages in the IMS access security framework 41
2 References 41
3 Definitions, symbols and abbreviations 42
3.1 Definitions 42
3.3 Abbreviations 43
Annex A: Enhancements to the access security for IP based services to enable NAT traversal for signaling messages 44
A.1 Scope 44
A.2 References 44
A.3 Definitions, symbols and abbreviations 44
A.4 Overview of the security architecture 44
A.5 Security features 44
A.6 Security mechanisms 45
A.6.1 Authentication and key agreement 45
A.6.2 Confidentiality mechanisms 45
A.6.3 Integrity mechanisms 45
A.6.4 Hiding mechanisms 45
A.6.5 CSCF interoperating with proxy located in a non-IMS network 46
A.7 Security association set-up procedure 46
A.7.1 Security association parameters 46
A.7.2 Set-up of security associations (successful case) 51
A.7.3 Error cases in the set-up of security associations 5554

A.7.3.1 Error cases related to IMS AKA 5554
A.7.3.2 Error cases related to the Security-Set-up 5554
A.7.3.2.1 Proposal unacceptable to P-CSCF 5554
A.7.3.2.2 Proposal unacceptable to UE 55
A.7.3.2.3 Failed consistency check of Security-Set-up lines at the P-CSCF 5655
A.7.3.2.4 Missing NAT traversal capabilities in the presence of a NAT 5655
A.7.4 Authenticated re-registration 5655
A.7.4.1 Void 5655
A.7.4.1a Management of security associations in the UE 5655
A.7.4.2 Void 5756
A.7.4.2a Management of security associations in the P-CSCF 5756
A.7.5 Rules for security association handling when the UE changes IP address 5857
A.8 ISIM 5857

Annex H (normative): The use of "Security Mechanism Agreement for SIP Sessions" [21] for security mode set-up 5958

Annex E: Improved IMS AKA for IPSec Traversal NAT 6059
E.1 Discussion 6059

Annex <X>: Change history 6362

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

This clause is optional. If it exists, it is always the second unnumbered clause.

1 Scope

Editor's Note: The current version of this document represents the state of discussion in 3GPP SA3 after meetings #39 and #40. It consists of a compilation of input documents to SA3 #39 and SA3 #40. The annexes describe solutions proposed for signalling protection for fixed broadband access to IMS. They are based on input documents to SA3 #39 and SA3 #40 (S3-050402, S3-050539, S3-050571). Section 5 (Analysis) represents the contents of the analysis contributions discussed (S3-050333, S3-050372, S3-050427, S3-050570). The current text in section 5 is not endorsed by SA3 at this point of time.

The scope of the present document is to study security requirements and solutions related fixed broadband access to IMS. Both solutions for ETSI TISPAN NGN R1 and ETSI TISPAN NGN R2 need to be studied. Based on this document, solutions to meet the fixed broadband access security needs are to be specified in TS 33.203 within the time frame of NGN R1 and NGN R2.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] S3-040990, "IMS security extensions", Ericsson, SA3#36 meeting

[2] S3-0401038, "BT Comments on S3-040990 IMS security extensions", BT Group, SA3#36 meeting

[3] 05TD161, "Feasibility of IPsec and TLS to provide SIP signalling security on the access in NGN/IMS", Ericsson and Alcatel, TISPAN#5 meeting

[4] "Datagram Transport Layer Security", Standard Track RFC candidate, http://www.ietf.org/internet-drafts/draft-rescorla-dtls-05.txt, Current state: RFC Editor Queue

[5] 05bTD078, "TLS based IMS access security architecture", Ericsson, TISPAN#5bis meeting

[6] S3-050239, "Scalability of IMS/TLS server certificate deployment", Ericsson, SA3#38 meeting

[7] S3-040720 "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", BT Group, SA3#35 meeting

[8] ECC Report 50: TECHNICAL ISSUES OF ESTABLISHING ANY-TO-ANY 2-WAY REAL-TIME COMMUNICATIONS OVER THE INTERNET, http://www.ero.dk/documentation/docs/doc98/official/pdf/ECCREP050.PDF

[9] S3-050048, BT Group

[10] S3-050242, Ericsson

[11] S3-050255, Siemens

[12] RFC 3948: UDP Encapsulation of IPsec ESP Packets

[13] RFC 3489: Simple Traversal of UDP Through Network Address Translators

[14] IETF Middlebox Communication (midcom) charter: http://www.ietf.org/html.charters/midcom-charter.html

[15] The Universal Plug and Play (UPnP) Forum : http://www.upnp.org/

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the [following] terms and definitions [given in ... and the following] apply.

example: text used to clarify abstract rules by applying them literally.

3.2 Symbols

For the purposes of the present document, the following symbols apply:

<symbol> <Explanation>

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

<ACRONYM> <Explanation>

4 Requirements

4.1 ETSI TISPAN NGN R1 Requirements

4.2 ETSI TISPAN NGN R2 Requirements

5 Analysis

5.1 IMS access security solution for NAT/FW traversal

Editor's Note: Section 5.1 is based on input documents S3-050372 and S3-050427 to SA3#39. The comments provided in S3-050427 to document S3-050327 are formatted as "List Bullet 2" in this section.

5.1.1 Introduction

This section discusses several alternatives for IMS access security solutions and highlights the pros and cons for each alternative when NAT/FW traversal is needed. The current IMS access security solution as specified in TS 33.203 is out

of the scope of the alternatives, as it doesn't work with NAT/FW devices. The early IMS Security solution specified in TR 33.978 doesn't provide SIP signalling protection on IMS level, thus this solution cannot be used in broadband access network as such.

The following IMS access security solution alternatives for NAT/FW traversal discussed in this contribution are:

- TLS
- IPSec tunnel mode (terminating to P-CSCF) with UDP encapsulation
- WLAN-IW scenario 3 with IMS access security (TS 33.203)

5.1.2 Alternatives

5.1.2.1 TLS

TLS has been already discussed in several earlier contributions seen in SA3, for example in [1]. The solution offers the following advantages:

- Provides privacy even for the first REGISTER message
  - not true if pskTLS with http digest aka for key establishment is used, as http digest aka has to be run before pskTLS can be set up.
- Availability of client implementation (part of IETF SIP standard)
  - true for TLS client, but not for pskTLS, DTLS or http digest aka client
- Mature and widely deployed mechanism
  - true for TLS client, but not for pskTLS, DTLS or http digest aka client
- Already very commonly deployed in fixed network environment
  - the problem with fixed networks is client authentication. Certificates are difficult to accept for many operators, fixed networks may have the problem that no ISIM /USIM may be available, passwords may not be acceptable as they may be copied.

The following disadvantages have been discussed in ([2], [3]):

- Does not solve media protection
- Cannot be used with UDP. However, by using Datagram TLS [4] a signalling message transported with UDP may also be protected
  - true, but there is no practical experience with DTLS
- TLS support need to be implemented in P-CSCF

The used authentication mechanism in TLS-based solution needs to be decided. Authentication based on TLS server certificates and HTTP Digest AKA is one option, another one is using PSK TLS. These options have been discussed in [5] and [6] but a detailed solution needs further work.

5.1.2.2 IPSec tunnel mode (terminating to P-CSCF) with UDP encapsulation

The proposal has been described in [7]. UDP encapsulation was proposed to be implemented with IMS AKA instead of IKE.

This solution has some advantages:

- Could provide also media protection

The following disadvantages of this solution need to be taken into account:

- IPSec implementation changes and UDP encapsulation termination is required in P-CSCF.
  - S3-050402 suggests that UDP encapsulation with tunnel mode may be provided as a modular add on to the Release 5 IPsec implementation,
- Need to refresh NA(P)T binding frequently
  - this is not an issue of IPsec vs TLS but rather one of TCP vs UDP, UDP needs refreshes more frequently, but the mobile environment needs to support UDP..
- RFC3948 (UDP encapsulation of IPSec ESP packets) states, that protocol assumes usage of IKEv1 or IKEv2
- If media is protected with the same IPSec tunnel, media flows go through P-CSCF. This is not an optimal solution from architecture or performance point of view.
- The IPSec implementation is IMS specific, which slows down adoption of solution in some of the terminal types used in broadband environment.

5.1.2.3 WLAN-IW scenario 3 with IMS Access Security (TS33.203)

The solution is based on existing 3GPP specifications in TS33.234 and TS33.203. Media is protected by tunnel mode IPSec between UE and PDG according WLAN-IW scenario 3. SIP signalling is integrity and optionally confidentially protected between UE and P-CSCF with transport mode IPSec inside the outer IPSec tunnel. Authentication is based in IMS AKA.

Another option is that the SIP signalling is protected only by tunnel mode IPSec to the PDG. In this case Network Domain Security is used between PDG and P-CSCF for signalling protection. This option still requires IMS level authentication to be used.

The above presented solution offers the following advantages:

- Provides media protection
- Based on 3GPP standardized mechanisms specified in TS 33.234 and TS 33.203
- Implementation support in 3GPP mobile terminals
- Flexible solution allowing to replace the inner IPSec with another solution

The first option of the solution has the following disadvantage:

From terminal point of view the performance is not optimal due to two IPSec connections

5.2 NAT device traversal and interoperability issues for IMS Rel-7

Editor's Note: Section 5.2 is based on input document S3-050333 to SA3 #39.

5.2.1 Introduction

Several alternative approaches have been presented to overcome problems related to NAT in IMS Rel-7. Before SA3 decides to fundamentally change Rel-5 IMS security, it should be clear which exact problems these approaches can overcome, and where they are still lacking. Changes to Rel-5 IMS should be minimal to avoid interoperability problems between Rel-5 and Rel-7 IMS.

5.2.2 Discussion

5.2.2.1 NAT-related difficulties

The problems with a NAT device at the UE site can be separated into different categories:

1a) Signalling protocol problems with NAT traversal

1b) Problems with incoming signalling connections

2a) Media protocol problems with NAT traversal

2b) Problems with incoming (protected?) media connections

SA3 already studied protocol-related NAT traversal issues 1a), but the problem of incoming connections deserves some more attention.

Problems related to incoming connections are explained in [8], for example. Without further measures, a NAT device does not allow incoming connections. During finalisation of Rel-5 IMS security, SA3 spent some time on specifying port handling in section 7.1 of TS 33.203. In those discussions, it was clarified that a P-CSCF must be able to establish a new signalling connection to the UE, despite the facts that the UE did already establish another TCP connection to the P-CSCF, and that TCP connections are bi-directional.

A similar problem exists with media connections: the Rel-5 IMS architecture allows direct media streams from one UE to another. An example shall show how potential solutions will affect the IMS architecture: To avoid the incoming connection, both UE could actively connect to a media gateway, which passes the data on. This would impose a change on IMS procedures and introduce a new network element. So even if implementations, architecture, protocols, and/or IMS procedures could be adapted to handle or work without incoming connections, such seemingly simple solutions will significantly deviate from IMS Rel-5 and break interoperability.

5.2.2.2 Solutions in the protocols

Different proposals to address NAT issues have been presented in SA3, e.g. [9], [10], [11]:

- Rel-5 IMS with UDP encapsulation [12]
- TLS
- Generic Access (ESP tunnel mode)

A fourth alternative is added which is directly comparable to TLS in the scope of this document. Problem 1a), which is solved whe using TLS, could also be addressed by "simply" (in terms of specification work) switching Rel-5 security from transport mode ESP to tunnel mode ESP:

Rel-5 IMS, using IPsec ESP in tunnel mode

STUN [13] is not considered a viable alternative, and is therefore not listed.

The following table lists the four proposals, with their pros and cons:

	Rel-5 IMS & UDP encaps	**Rel-5 IMS with ESP tunnel mode**	**TLS**	**Generic Access (GA)**
Incoming signalling connections	yes	no	no	see below *)
Media protection and incoming media connections	to be defined	to be defined	to be defined	protection included, incoming connections see below *)
UDP support	yes	yes	with datagram TLS?	yes
Double encryption	no	no	no	yes, for signalling
Rel-5 compatibility	high	low	low	high (but complex add-on)

*) Incoming connections with GA could be allowed by using UDP encapsulation.

The row "Rel-5 compatibility" deserves some explanation, which is given in section 4.

Conclusion: whereas all proposals address problem 1a), only UDP encapsulation provides support for incoming connections 1b).

5.2.2.3 Solutions in the NAT device

For completeness' sake it should be mentioned that there are some means to configure incoming connections to specific ports in most NAT devices. These connections are then redirected to specific addresses behind the NAT device. Many NAT devices support manual and static configuration of this port redirection, and there are also interfaces to allow dynamic reconfiguration ([14], [15]). These NAT-device based means do not seem suitable for Rel-7 IMS due to several reasons:

- They make the IMS solution dependent on features of the NAT device
- They may open security holes in the IMS, because the device control interface specification is out of scope for 3GPP
- Not all are (completely) standardised

5.2.2.4 Multiple clients behind one NAT device

It should be clarified by TISPAN or SA1, if there is a service requirement to support multiple UE behind one NAT device at the same time, or if such a requirement is envisaged for a later release. This must be taken into account when selecting the solution. None of the NAT traversal protocol solutions mentioned in section 2.2 is capable of supporting incoming connections for multiple clients without modifications.

In addition to the incoming connections problem, new trust and charging relevant questions arise:

- Does the owner of the NAT device decide who can access IMS services through the device? How?
- How is bearer charging involved?
- Interoperability of IMS Rel-7 and Rel-5

It is assumed that a NAT traversal solution which is independent from Rel-5 IMS can be added in a modular way. This will provide less compatibility problems than a change in the IMS mechanisms or the IMS architecture itself. From an IMS perspective, UDP encapsulation and GA can be seen as "bearer level", and direct impact on Rel-5 IMS security will be small.

OMA specifications rely on the existing Rel-5 IMS specifications, which have been stable for some time. Will OMA quickly adopt a Rel-7 IMS when it is very different from Rel-5? If 3GPP now significantly changes IMS protocols in Rel-7, there will be interoperability problems. We could end up with three incompatible security solutions:

- early IMS, defined because Rel-5 IMS is not implemented yet
- Rel-5 IMS, with products currently being implemented
- Rel-7 IMS

In that case it could be considered to drop Rel-5 IMS security completely and live with only two incompatible solutions.

5.3 Analysis of Enhanced 3GPP R5/6 Access Security Mechanism

Editor's Note: Section 5.2 is based on input document S3-050570 to SA3 #40.

5.3.1 Introduction

The following sections analyse the Enhanced 3GPP R5/6 Access Security Mechanism (E3G-ASM) that is presented in [S3-050402]. Section 5.3.2 comprises from sub-sections that present detailed analysis of specific aspects of E3G-ASM. The special focus is given to issues related to UDP encapsulation [RFC3948].

5.3.2 Analysis

When dealing with transport security, it can be noted that NAT traversal in general becomes more transparent to the security layer, the higher up in the stack you apply the security. A session/application layer security solution (like TLS) will not need to bind the security association to the transport addresses and ports, but can do the mapping on other identities. In contrast, the lower in the layers you apply the security (such as IPsec), the more "fixes" must be done to the solution to be able to handle the NATs. While IPsec partly was created to be "transparent" to the application, the E3G-ASM solution has proven that the transparency is all long gone. In fact, not only is the application involved in the key management, but will also need to handle the NAT traversal. This creates a very tight coupling between the network layer and the application, which in many respects can be questioned if it is desirable (this problem is sometimes referred to as layer violation).

This analysis focus on four aspects of the E3G-ASM:

1) Implementation (and upgrades)
2) Co-existence with other applications
3) Immaturity of Standard
4) NAT compatibility problems

5.3.2.1 Implementation

The E3G-ASM solution is often regarded as a relatively simple extension to the existing access security mechanism of 3GPP R5/6. However, implementation wise, it is not always that straightforward. The Figure 1 shows the coarse diagram from typical, existing IPsec implementation. Kernel needs to have a support for IPsec, and in needs to include a standard base Application Programming Interface (API). User space has to have some kind of keying daemon and policy database. Numbers in figures indicate the order of events. The events in Figure 1 are the following:

1) Keying daemon reads the policies from database.
2) Keying daemon registers itself to the kernel.
3) Application tries to send information to the network.
4) The SA decisions part of the kernel informs keying daemon that the required SA does not exist.
5) Keying daemon initiates a key exchange procedure (e.g. IKE).
6) Keying daemon creates a SA in the kernel.
7) The data sent by the application is passed by the IPsec implementation in the kernel. After IPsec procedures the data is sent to the network.

Today, there are only few IPsec implementations that support UDP encapsulation (and most likely that current TS 33.203 implementation does not have UDP encapsulation support), so, therefore, a typical IPsec implementation without UDP encapsulation is chosen as a reference. IPsec implementations are operating system specific. IPsec implementations on some operating systems may slightly differ from the diagram presented in Figure 1. However, the same principles apply to most of them.

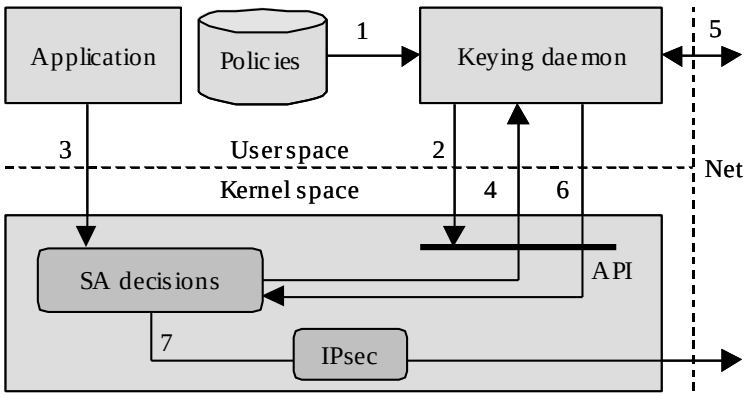


Figure 1: Coarse diagram from typical, existing IPsec implementation

The IPsec implementation required by the E3G-ASM is presented in the Figure 2. The events are slightly different when compared to the typical IPsec implementation, and the biggest changes are emphasized with bold font:

1) Keying daemon reads the policies from database.
2) Keying daemon registers itself to the kernel.
3) Application tries to send information to the network.
4) The SA decisions part of the kernel informs keying daemon that the required SA does not exist.
5) Keying daemon initiates a key exchange procedure, **which is done on SIP layer.**
6) Keying daemon creates a SA in the kernel. **This procedure call also needs to convey the information that UDP encapsulation is needed.**
7) The data sent by the application is passed by the IPsec implementation in the kernel. **SA decisions part of the kernel needs to be able to make a decision whether or not the data needs to be UDP encapsulated.**
8) IPsec packets are UDP encapsulated, and then sent to the network.

Part where big changes are needed are displayed with grey, diagonal pattern. Hardest things to implement are the three different changes to the kernel. First change is the API. The API needs to be modified in such a way that it is possible to convey UDP encapsulation information from user space to kernel space. The second change is that the part making the SA decisions in the kernel needs to be able to determine whether the application data needs to be UDP encapsulated or not. From implementer's point of view this is not very easy, because it is possible that a new routing decision is needed, and there might also be a need to create a new queue for the data packets. The third change is the insertion of UDP encapsulation functionality itself, which is not present in the most current IPsec implementations. In case some or all of these are implemented in hardware, changing them is especially difficult if not even impossible.

Also the user space requires a new keying daemon, which does the key exchange on SIP layer. It is good to keep in mind that these changes proposed by E3G-ASM are not only related to the terminals, put also to the system software in P-CSCF needs to be changed.

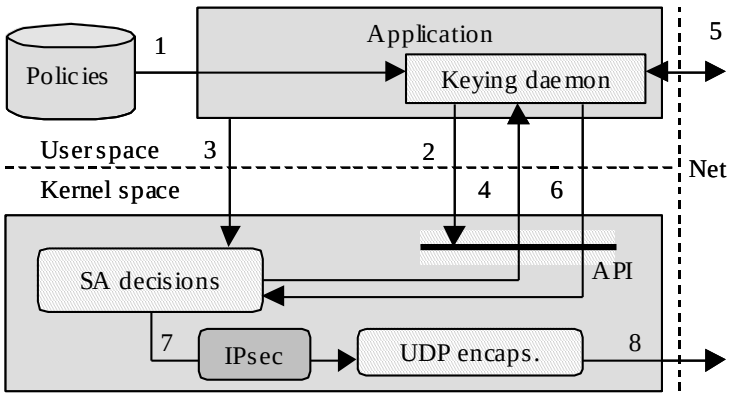


Figure 2: Coarse diagram from IPsec implementation done according the E3G-ASM

If terminals are based on some closed source operating system, such as MS Windows or Macintosh, it is impossible for a 3rd party to make changes to the kernel. Instead, special plug-ins must be developed (in the worst case a full IPsec implementation, which may conflict with the existing one that is used). This fact will probably slow down the adoption of NGN terminal software among customers, because there will not be many implementations to choose from in the market. It is also noteworthy that the NAT traversal problem is present in fixed access networks, and the vast majority of terminal there are using closed source operating systems.

It should also be noted that hardware implementations are sometimes used for servers and specialized terminals. Adding UDP encapsulation in this case is not non-trivial but will create a large cost.

5.3.2.2 Co-existence of Applications

A potential problem for E3G-ASM is the co-existence of other application utilizing the IPsec engine (such as Virtual Private Network (VPN) applications or other applications). These applications typically utilize particular standard keying daemon (such as IKE/IKEv2) or may in different (non-standard) keying deamon. There is a general problem of using more than one keying daemon as there may become conflicts while setting up policies and handling SPI and SAs.

Conflicting policies could easily be created by e.g., a corporate VPN client and a IMS application where the IMS application first set up a policy between UE and P-CSCF, which then the corporate VPN client will try to override (there may even be conflicting policies for I-WLAN clients and IMS applications, in case the IMS application sets up a policy before the I-WLAN client is activated). Another problem is of course SPI handling. When more than one keying deamon competes of the same SPI space, conflicting SPI's may appear as a result.

5.3.2.3 Immature standards

UDP Encapsulation [RFC3948] is a new technique, and the first implementations are just starting to emerge. Some problems have been spotted on this technique. The most notable problem is the poor interoperability in the situations where NAT device between two UDP encapsulation capable devices suddenly reboots. Current specification [RFC3948] does not give enough details on how this kind of situation should be handled. As a consequence, it is possible that UDP encapsulation capable devices from vendor A do not work with the devices from vendor B.

Another problem with E3G-ASM is that it requires changes to the IANA registry, which is created by [RFC3329]. Seemingly, a Standards Track IETF RFC or a separate IAB decision is needed in order to make that change. Arguably, it might take too long for such a process, considering the schedule of NGN R1.

5.3.2.4 Poor compatibility with NATs

E3G-ASM sets up two Security Associations (SAs), and it uses UDP Encapsulation. These two mechanisms inflict some problems to NAT devices. The use of UDP as a transport protocol mandates frequent keep-alive messages. The following is a citation from "IP Network Address Translator (NAT) Terminology and Considerations" [RFC2663]:

"Many heuristic approaches are used to terminate sessions. You can make the assumption that TCP sessions that have not been used for say, 24 hours, and non-TCP sessions that have not been used for a couple of minutes, are terminated."

It is obvious that not all NAT implementations are in direct accordance with this text, but is gives a pretty good estimate from the scale of NAT binding lifetimes. Real life example: UDP bindings in the NAT implementation of Linux 2.6.10 timeout in 60 seconds, when compared to TCP bindings which timeout in 5 days.

The fact that E3G-ASM uses two SAs causes the need to make two bindings to NAT devices per one signalling connection. This is not a problem to such NAT devices that reside on customer premises, but it might be a problem for those NAT devices that reside on the premises of access network provider. The maximum number of bindings per outer IP address in NAT devices is 65536.

In addition, IPsec itself has some unsolved problems with NATs. Conflicted situations are possible, and in some scenarios even probable, when using either tunnel or transport mode. These conflicted situations are explained in detail in Section 5 of [RFC3948].

5.4 Comparison of solutions for IMS signalling protection that are proposed to be added in 3GPP Release 7 (SA3 drafting Group at SA3 #41)

Editor's note: This section incorporates S3-050827, which was a result of the SA3 drafting group at SA3 #41. This constitutes a summary of views presented during the SA3 meeting, but no endorsement by SA3 was achieved.

5.4.1 Introduction

This document gives a comparison of the two solutions that are proposed to be added for IMS signaling protection in 3GPP Release 7:

1. TLS-based access security: The latest draft CR is available as S3-050762. In this CR the network-side endpoint of the TLS tunnel is verified using an IMS AKA related key. However, at SA3#41 it was instead proposed to use "regular" TLS for server-authentication and client authentication based on Digest AKA. It is this new proposal that is the basis for the comparison in this document. Note that an alternative variant based on PSK-TLS was also discussed, but it is no longer under consideration.

2. IPsec-based access security with UDP encapsulation: The latest draft CR is available as S3-050533.

The comparison is based on the arguments presented in the earlier contributions in SA3, and on the discussions during SA3#41, and are categorized in order to help the decision making in SA3. Since, from security point of view, there is no significant difference between the two solutions, SA3 needs to make the decision based on the other aspects. In this document we concentrate on the arguments that should affect the decision.

5.4.2 Comparison

Aspect	TLS based access security	IPsec based access security
Availability	▪ Regular TLS is already available in many SIP client implementations. This makes the deployment of TLS cheap and quick. Furthermore, 3GPP Release 6 UE already supports TLS for presence service. However, the combination of TLS and Digest AKA is specific to IMS deployment. ▪ Mature and widely deployed mechanism.	▪ The IPsec implementation is IMS specific, which may slow down adoption of solution in some of the terminal types used in broadband environment. ▪ There are only a few IPsec implementations currently existing that support UDP encapsulation according to the RFC. However, most VPN implementations are converging towards the RFC. All major VPN vendors support NAT-traversal using UDP encapsulation in their gateways and VPN clients in a proprietary way. ▪ 3GPP Release 6 UE already supports IPsec transport mode for IMS. 3GPP Release 6

		UE also supports IPsec with UDP encapsulation for WLAN scenario 3 and for generic access to A/Gb, although the key management is different.
Interoperability	▪ TLS interoperates easily with all existing PC applications, and does not require changes to the operating system	▪ UDP Encapsulation is a new technique, and the first implementations are just starting to emerge. Some problems have been spotted on this technique. One example of the problems is the poor interoperability in the situations where NAT device between two UDP encapsulation capable devices suddenly reboots. However, there are provisions in RFC3947, section 7, for the case of NAT rebooting also for the case where IKE is not used. Even if they did not suffice, it is questionable whether a rebooting NAT in a telecom environment is a frequent situation. Anyway a client will try to re-register eventually when communication attempts fail.
Complexity	▪ Alignment with IETF SIP standard, except for the use of Digest AKA. ▪ TLS is easy to integrate to application layer. Software development will be easier and faster. ▪ For interoperability with Release 5/6 IMS, both terminal and network may need to support both IPsec and TLS based solutions for IMS security.	▪ For interoperability with Release 5/6 IMS, both terminal and network may need to support both IPsec transport mode and IPsec with UDP encapsulation for IMS security. The burden to support both IPsec transport mode and IPsec UDP encapsulation, is less than the burden to support IPsec and TLS based solutions. ▪ A potential problem is the co-existence of other applications utilizing the IPsec engine (such as Virtual Private Network (VPN) applications or other applications). This is also a potential problem with Release 5 IMS and 3GPP IP access in 3G-WLAN interworking. The extent of this problem depends on the particular implementation. ▪ According to RFC3947 (Negotiation of NAT-Traversal in the IKE) and RFC3948 (UDP encapsulation of IPsec ESP packets), UDP encapsulated packets for ESP and IKE must use the same well know port 4500. They are distinguished by a payload starting with either four zero octets (IKE) or a different value (the ESP SPI). Thus, in standard scenarios, an IKE NAT-T capable daemon listens on port 4500 and demultiplexes IKE and ESP traffic. In this configuration there may be implications for the implementation in case IKE is also used on the same network interface of a P-CSCF, since the standard assumes the same port number for UDP encapsulated IKE and ESP traffic. Note that there is no problem on the client side, as an IKE daemon can perfectly co-exist with an IMS based IPSec usage, including NAT-T. On the P-CSCF there might be an issue, but this strongly depends on the concrete implementation of IKE and

		IPSec. Nevertheless, it can be considered to be quite unlikely that a P-CSCF will run an IKE daemon on port 4500 on the Gm interface. The IMS Gm interface is surely a "full time job" and the same physical interface is probably not used for other purposes. It goes without saying that other physical interfaces on a P-CSCF (e.g. used for OAM), of course, can definitely run an IKE daemon on port 4500.
Other	▪ If the P-CSCF is in a visited network different to the home network then cross certification between the involved operators is required. ▪ TLS cannot be used with UDP. However, by using Datagram TLS a signalling message transported with UDP may also be protected. Although DTLS is in RFC editor's queue as standards track and supported in OpenSSL, we cannot assume DTLS to be widely available for the near future. ▪ No issues when multiple clients are used behind a NAT. ▪ A TLS connection must be kept alive for an extended period of time in order to guarantee NAT traversal. This may raise implementation issues at the client and scalability issues on a P-CSCF. ▪ A TLS-based solution will suffer from a NAT rebooting as UE terminating messages will not reach their destination until the client takes action to reinitiate TLS. ▪ Any TLS solution needs additional roundtrips for the setup of the TLS tunnel. This leads to a higher delay for IMS registration. ▪ The adoption of TLS for IMS extensions may facilitate convergence with IETF SIP standard. ▪ Authenticated re-registration and change of TLS keys needs further study. Furthermore, the correct version of HTTP Digest AKA to use needs further study.	▪ RFC3948 assumes usage of IKEv1 or IKEv2. However, it does not exclude the usage of UDP encapsulation in the context of other IPSec negotiation mechanisms. In S3-050402 and S3-050533 it was shown that this is indeed feasible. ▪ The P-CSCF shall not accept registration attempts from UEs with the same address and protected server port in order to avoid ambiguities. Such situations may occur in case of multiple UEs behind the same NAT, which are assigned the same public IP address by the NAT. In S3-050402 and S3-050533 considerations were given how to cope with such a situation. In addition, one might consider to take measures to almost entirely prevent an accidental clash of protected server ports. For example, one could ask a UE to determine the protected ports by performing a hash calculation that includes its private IP address or other individual information.

5.5 Signalling Protection

5.6 Media Protection

6 Conclusions

:

Annex A:
Approaches to TLS based IMS security solutions

Editor's Note: Annex A is based on input document S3-050571 to SA3#40.

A.1 Introduction

In SA3#39 TLS based IMS access security solutions were proposed in [S3-050407]. One TLS solution was based on traditional TLS [RC2246] (i.e. certificate based server authentication and shared key based client authentication) and the other solution was based on PSK TLS [PSKTLS]. This contribution further elaborates on both of the proposals.

A.2 Problem statement

One of the challenges of TLS based access security in IMS is the roaming case, in which the UE establishes the TLS tunnel to the P-CSCF in the visited network. In the roaming case the UE needs to be able to trust the TLS tunnel, i.e. the UE needs to know that an authorized entity and not a man-in-the-middle (MitM, that could convey the IMS signalling) is on the other end of the TLS tunnel.

In case of PSK TLS, the trust to the TLS tunnel is self-evident due to the use of the session keys CK/IK in setting up the TLS tunnel. MitM does not have access to CK/IK.

In case of a certificate based TLS, the issue is not so straightforward, since the UE should be able to trust the server side certificate. This trust trust would require either cross-certification between operators (PGP model) or globally trusted Certificate Authorities (CA) as part of Public Key Infrastructure.

It has been noted in [S3-050239] that cross-certification may have scalability problems. In addition, certificates revocation may be a problem for certificates in general.

The following chapter introduces an IMS AKA asserted solution to the certificate trust problem, where P-CSCF and UE bind session keys CK and IK to the TLS tunnel and provides the assertion of the server side certificate.

It should be noted that when only "TLS" is mentioned in the present document then both certificate based TLS (I-TLS) and PSK TLS are meant. Otherwise the TLS modes are mentioned explicitly, if the mode is not clear from the context.

A.3 Solution

A.3.1 IMS AKA for authentication

IMS AKA is the mechanism in IMS Rel-5/6 [33203] for the mutual authentication between the UE and the S-CSCF in the home network. Regardless of the signalling access security solution between UE and the P-CSCF in Rel-7, it is assumed that IMS AKA will be used on top of the access security solution.

It should be noted that both TLS based solutions (i.e. I-TLS and PSK TLS) presented in the following chapters are independent of the chosen AKA version, i.e. AKAv1 or AKAv2.

A.3.2 TLS protection

TLS provides transport-layer security, (i.e. data integrity, data origin authentication, data confidentiality and protection against message replay) over connection-oriented protocols. For the purposes of this document, TCP can be specified as the desired transport protocol within a "Via" header field value or a SIP-URI.

TLS is well suited to architectures in which hop-by-hop security is required between hosts with no pre-existing trust association.

A.3.3 IMS AKA asserted TLS (I-TLS) signalling protection solution

This chapter introduces an IMS AKA assertion solution to the certificate trust problem, where the P-CSCF and UE bind session keys CK/IK to the TLS tunnel and provides the assertion of the server side certificate. In order for the UE to be able to trust the server side certificate, the P-CSCF calculates a MAC over the server side certificate with CK/IK that P-CSCF has received from the S-CSCF and sends this to the UE. By verifying this MAC (called server token in this contribution) the UE is able to trust the server side certificate and the corresponding TLS tunnel. The UE in turn calculates a MAC over the server token using CK/IK, and sends this to the P-CSCF. By sending this MAC (called client token in this contribution) the UE acknowledges that it received and accepted the server token.

It should be noted that the server side certificate used by P-CSCF does not need to be part of any particular PKI for the user to trust it and it can be a self-signed certificate, if the mechanism described in this contribution is used. The only requirement on the certificates is that they are formed according to the general format and that the public key of the server is included properly. The client will not need to verify the CA signature (as this verification is replaced by the server token).

The UE-side verification of the server token is not intended necessary for protecting against a client-impersonation and MitM session hijacking attacks because the server will notice that the "client token" is wrong and abort the procedure. But if the client might send some confidential data to P-CSCF at the end of the procedure, then it is necessary for the client to explicitly authenticate P-CSCF. UE-side authentication is intended for this.

A.3.3.1 Overview of IMS AKA asserted TLS solution

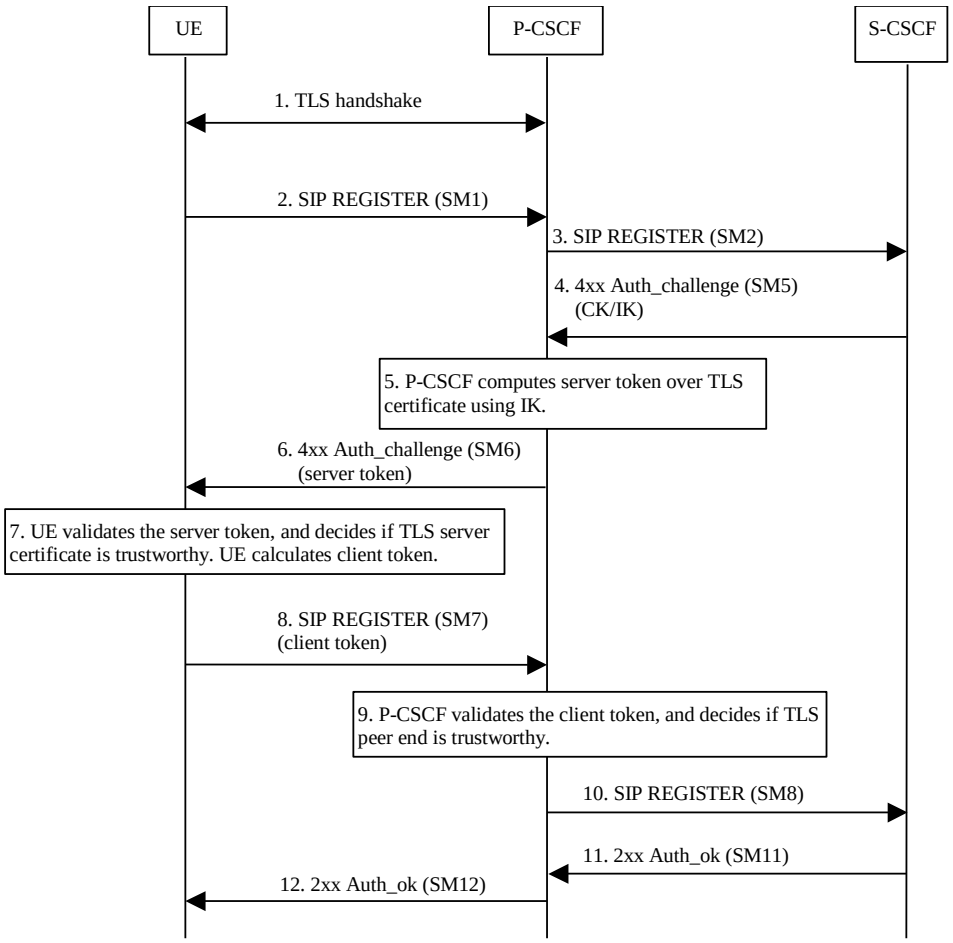


Figure 1: IMS AKA asserted TLS based IMS access security

The IMS AKA asserted TLS based IMS signalling protection solution is depicted in Figure 1. It should be noted that the IMS registration messages are the same as in IMS Rel-5/6 with the exception that the server token is carried in message SM6 and client token is sent in message SM7.

The procedure is as follows:

1. UE and P-CSCF perform full TLS handshake. The P-CSCF uses server side certificate for the TLS tunnel. The UE authenticates the P-CSCF at TLS level by using the server certificate provided by the server. To avoid unnecessary computations (and possible user interaction), the UE need not verify the CA signature in the certificate, as it can simply accept the certificate. At this stage the UE will not be sure that it can trust the provided certificate and the corresponding TLS tunnel.

2. UE starts IMS registration procedure with SM1 message. This message is the same as in IMS Rel-5/6.

3. P-CSCF relays IMS registration message in SM2 message. This message is the same as in IMS Rel-5/6.

4. S-CSCF sends the authentication challenge with CK/IK. This message is the same as in IMS Rel-5/6.

5. P-CSCF strips off CK/IK from the authentication challenge as in IMS Rel-5/6. In addition P- CSCF calculates the server token (i.e. MAC) over the server certificate using IK, and it appends the server token to the challenge message.

6. P-CSCF sends the authentication challenge to the UE in SM6. This message is the same as in IMS Rel-5/6, except that it carries also the server token.

7. UE processes the authentication challenge message as in Rel-5/6, e.g. it computes the session keys CK and IK. In addition the UE uses IK to validate the server token, i.e. it calculates a MAC over the server certificate of the TLS tunnel. If the computed MAC equals with the MAC received in the authentication challenge, the UE is able to trust the TLS tunnel. Note that the MAC over the certificate will give a guarantee that the P-CSCF is trusted by the home

network (If P-CSCF is not trusted by the home network it will not have access to IK). If the MAC verification fails, the procedure is aborted. Otherwise, the UE then calculates the authentication response. In addition, the UE calculates an authorization verification token (client token) to acknowledge that it received and accepted the server token, which is a MAC computed over the server token using IK.

8. UE sends the authentication response and client token to the P-CSCF in message SM7.

9. P-CSCF strips off and validates the client token. The client token is verified by the P-CSCF by calculating a MAC over the same field as the UE did, and then comparing the outcome with the client token. If the verification fails, the procedure is aborted.

10. P-CSCF relays IMS registration message in SM8 message. This message is the same as in IMS Rel-5/6.

11. If the user has been successfully authenticated, the S CSCF sends a 2xx Auth_OK message to the P CSCF in message SM11. This message is the same as in IMS Rel-5/6.

12. P CSCF forwards the 2xx Auth_OK towards the UE in SM12. This message is the same as in IMS Rel-5/6.

A.3.3.2 Interoperability with IMS Rel-5/6

A.3.3.2.1 Using security agreement

The UE and P-CSCF negotiate the security mechanism using Security agreement (Sec-agree) negotiation, which is specified in RFC 3329 [RFC3329]. If the UE supports TLS, it may start the communication with either with TLS or Sec-agree.

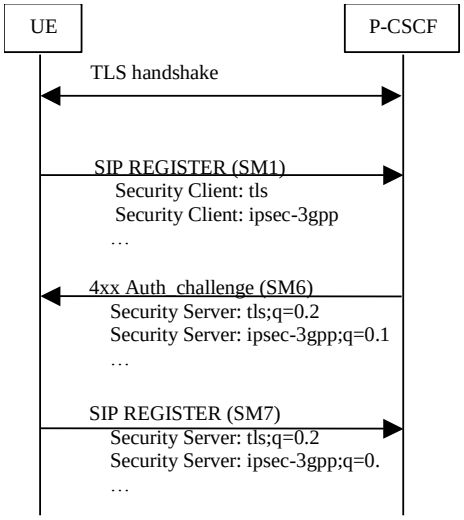


Figure 2: TLS is set up in the beginning

UE starts with TLS handshake and the Sec-agree negotiation is run in the following messages to confirm the choice of the security mechanism. Starting with TLS handshake has the benefit that the negotiation is protected from message SM1 and it does not add any roundtrips to the flow in the normal case. The Sec-agree negotiation does not impact the established TLS session if

1. TLS was the only mechanism supported by the UE and/or

2. TLS was the P-CSCF preferred mechanism (This case is shown in figure above).

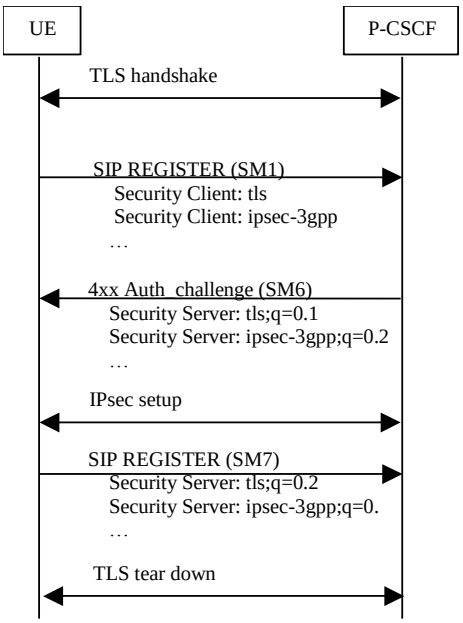


Figure 3: IPsec is setup due to P-CSCF preference

However, in the case 2 above, if for some reason the P-CSCF preferred mechanism is not TLS and also the UE supports this other mechanism (i.e. IPsec), then the preferred mechanism is taken into use and TLS tunnel is disconnected after the new mechanism is set up, see figure 3. In this case the message SM7 is transferred over the IPsec connection. The benefit of using TLS in this case is that the negotiation is protected from its beginning.

The UE may also start with Sec-agree before the TLS tunnel is set up. This is described in the Figure 4 below.

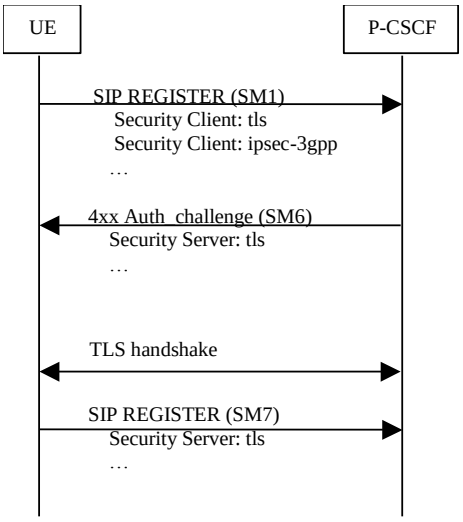


Figure 4: TLS is set up after Sec-agree

The UE indicates IPsec and TLS in Sec-agree, but the P-CSCF supports only TLS, therefore TLS is chosen.

A.3.3.2.2 Fallback to Rel-5/6

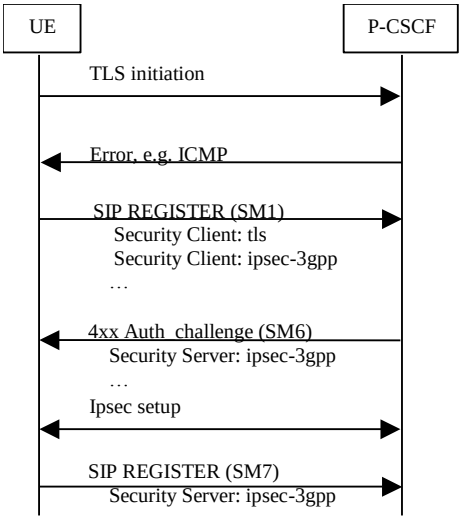


Figure 5: Fallback to IPsec.

This case presents the fallback to Rel-5/6 IPsec. UE starts with TLS handshake, which is rejected by the P-CSCF since it supports only Rel-5/6 IPsec. When receiving the error message the UE falls back to Sec-agree. Then the UE and P-CSCF negotiate the use of IPsec as in Rel-5/6.

It should be noted that since the error message from the P-CSCF cannot be authenticated by the UE, i.e. it could be sent by an attacker, the following Sec-agree negotiation may still lead to establishment of TLS. This is of course possible if both UE and P-CSCF support TLS.

Also in this case, the UE may also try to originally negotiate the security options before initiating the TLS tunnel set up.

It should also be noted that the case where the UE supports only IPsec and P-CSCF supports both TLS and IPsec is not described here since the situation is similar to the current Rel-5/6 solution.

A.3.3.3 The details of the token

As seen in Figure 1, the server and client tokens are carried in messages SM6 and SM7. The following shows an example how the server token and client token could be created and transported.

The server token (s_token) consists of a MAC value that is calculated over the server side certificate using HMAC-SHA1-96 [RFC2404] as algorithm and IK as the key.

The resulting MAC value is included as a parameter in the WWW-Authenticate header of 4xx Auth_challenge message (SM6) in the similar way as the IK and CK are transported from the S-CSCF to P-CSCF in corresponding WWW-Authenticate header of 4xx Auth_challenge message (SM5).

The client token (c_token) is a MAC that is calculated over the server token using HMAC-SHA1-96 as algorithm and IK as the key.

The client token is carried in the Authorization header of the authenticated REGISTER message (SM7).

An alternative way of calculating the tokens would be to use all available key material CK/IK and the generic key derivation function recommended by SAGE, and described in TS33.220

Similarly to the transport of CK and IK in Rel5/6 IMS, the transport of s_token and c_token within Digest headers is to be specified in TS 33.203 [33203], i.e. an internet draft is not needed.

An example of the WWW-Authenticate and Authorization headers carried in messages SM6 and SM7 is given below.

SIP/2.0 401 Unauthorized

…

WWW-Authenticate: Digest realm="registrar.home1.net", nonce=base64(RAND + AUTN + server specific data), algorithm=AKAv1-MD5, s_token="00112233445566778899aabb"

…

REGISTER sip:registrar.home1.net SIP/2.0

...

Authorization: Digest username="user1_private@home1.net", realm="registrar.home1.net", nonce=base64(RAND + AUTN + server specific data), algorithm=AKAv1-MD5, uri="sip:registrar.home1.net", response="6629fae49393a05397450978507c4ef1", c_token="ffeeddccbbaa112233445566"

…

If the UE does not support TLS, the s_token and c_token fields shall not be included by the P-CSCF in SM6 and SM7 messages

A.3.4 PSK TLS for signalling protection

This chapter describes how pre-shared key (PSK) TLS [PSKTLS] is used for IMS signaling protection. In IMS signalling protection context, PSK TLS has two very important benefits if compared to "normal" TLS (i.e. based on server side TLS certificates, and SIP Digest based client authentication):

· PSK TLS is easier to deploy securely. In "normal" TLS, we need to worry a lot about root CA's, certificate revocations, cross certification, and MitM attacks. With PSK TLS, all these problems disappear. It should be noted that the TLS solution presented in chapter 3.2 also overcomes these problems.

· PSK TLS works more easily for both directions. In "normal" TLS, we need to open the TLS session with SIP registration, and leave the TLS session open for all subsequent communication. There is no way for SIP proxy (P-CSCF) to open TLS to the client. With PSK TLS, the P-CSCF is able to open the TLS connection. It should be noted that the TLS solution presented in chapter 3.2 also overcomes these problems. In the presence of NAT this however requires that the TCP connection is left open.

Figure 6 demonstrates the solution details.

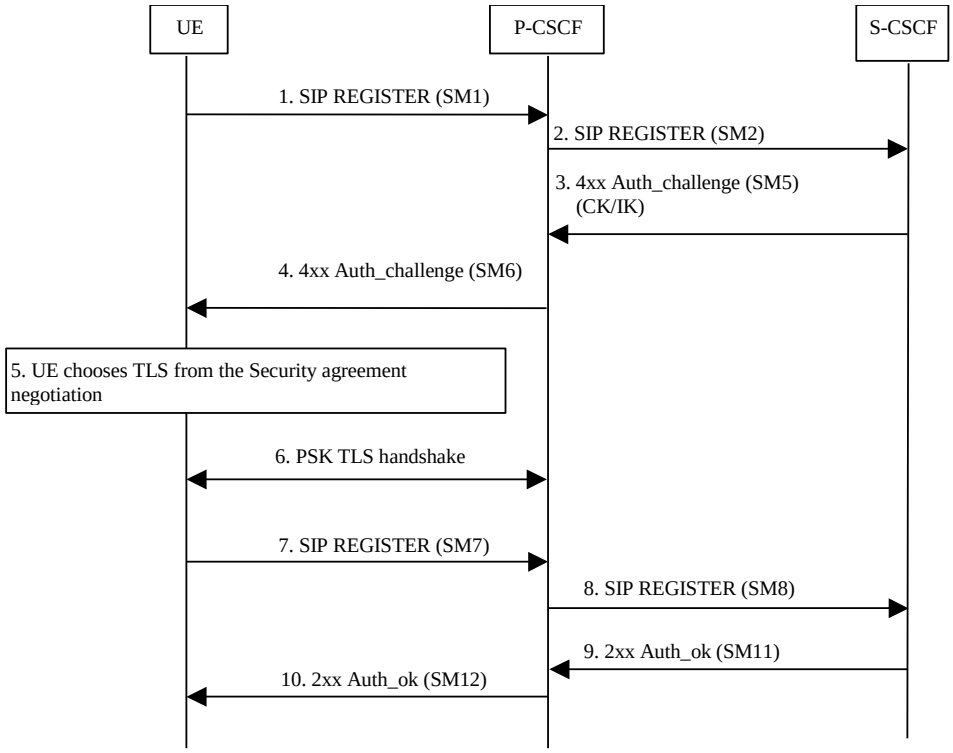


Figure 6: PSK TLS based IMS access security

Protocol details are as follows:

1. UE starts IMS registration procedure. The UE indicates TLS as an alternative security mechanism in "SIP security agreement" [RFC 3329].

2. P-CSCF relays IMS registration message in SM2 message. This message is the same as in IMS Rel-5/6.

3. S-CSCF sends the authentication challenge with CK/IK. This message is the same as in IMS Rel-5/6.

4. After removing the session keys from the response, P-CSCF forwards the response to the UE. This message is the same as in IMS Rel-5/6.

5. The UE follows the rules of RFC 3329, and chooses TLS as the security mechanism.

6. The UE and P-CSCF agree to use PSK TLS during the TLS handshake.

7. The UE continues with normal IMS registration procedure. UE sends the authentication response to the P-CSCF in message SM7. This message is the same as in IMS Rel-5/6.

8. P-CSCF relays IMS registration message in SM8 message. This message is the same as in IMS Rel-5/6.

9. If the user has been successfully authenticated, the S CSCF sends a 2xx Auth_OK message to the P CSCF in message SM11. This message is the same as in IMS Rel-5/6.

10. P CSCF forwards the 2xx Auth_OK towards the UE in SM12. This message is the same as in IMS Rel-5/6.

The interoperability towards Rel-5/6 is straightforward since it can be negotiated via the Sec-agree.

Presented solution with PSK TLS corresponds to the security level of the current IMS signalling protection. This means that initial registration message, and some error messages cannot be protected between UE and P-CSCF. PSK TLS is not currently among the security mechanisms of RFC 3329. However, this is not needed since the "tls" parameter can be used in this case, and the TLS cipher suits can be negotiated within TLS handshake.

A.4 Discussion

TCP state in P-CSCF

When a NAT is present between the UE and P-CSCF, the TCP connection below the TLS layer needs to be left open to enable continuous communication and communication that is initiated from the P-CSCF. This means that the P-CSCF needs to keep TCP state information for the UEs that the P-CSCF is communicating with.

It should be noted that for IPsec implementation, similar problems will arise, where a TCP connection (when used at least once) must not be closed down as this would result in problems if a new TCP connections will be needed within a short period of time. The problem is that 1) the IPsec SA is based on a specific client/server port pair, and 2) if a TCP connection is closed down, it will take some time before the resources for the connection is released fully and a new TCP connection can be established using that particular port pair. The only current viable solution for this currently is to keep the TCP connection open as soon as a TCP connection has opened. As current non-compressed SIP messages often exceed the 1400 bytes (the limit for UDP), it is most likely that a TCP connection will be needed to be kept also for IPsec based solutions. It is not clear if signalling compression will be used in TISPAN. Moreover, in this case the TCP connection cannot be established on demand, but it needs to be setup in the beginning of communication since it is setup with the Sec-agree negotiation.

IPsec implementations need to keep IPsec SA information in addition to the TCP state information as described above. In particular, the TCP state information can be compared to the state information needed for keeping IPsec connections for each user in P-CSCF. In a common operating system setup KAME/FreeBSD the state information needed by a TCP connection, i.e. the size of the so called TCP control block (excluding the buffers), was measured to be 328 bytes and the state information needed by IPsec (including two pairs of SAs [560 bytes] and one Security Policy entry per SA [576 bytes] plus SA index header [132 bytes]) was measured to be 1268 bytes. This shows that the ratio is 1268/328 i.e. approximately 3,8 times for the benefit of TCP. Therefore, the conclusion is that the resources needed for the TCP state information are not regarded to be an issue compared to the other information the P-CSCF needs to hold. (The state information needed in Kernel is compared here since it is assumed to be more difficult to optimize and the storage is more limited than in application level implementations.)

Incoming signalling connections

In TD S3-050333 [S3-050333] it was stated that "During finalisation of Rel-5 IMS security, SA3 spent some time on specifying port handling in section 7.1 of TS 33.203. In those discussions, it was clarified that a P-CSCF must be able to establish a new signalling connection to the UE, despite the facts that the UE did already establish another TCP connection to the P-CSCF, and that TCP connections are bi-directional." However, neither SIP RFC 3261 [RFC3261] nor section 7.1 of TS 33.203v6.7.0 reflects this understanding. Instead section 7.1 has two notes that say that existing TCP connections may be re-used: "Both the UE and the P CSCF may set up a TCP connection from their client port to the other end's server port on demand. An already existing TCP connection may be reused by both the P CSCF or the UE; but it is not mandatory." Therefore, it is assumed that no such requirement exists that was mentioned in S3-050333. In fact, IETF SIP WG is currently working on this specific issue, see [SIPDRAFT]. The approach in SIP WG is described in the introduction of the draft:

"The key idea of this specification is that when a UA sends a REGISTER request, the proxy can later use this same connection to forward any requests that need to go to this UA."

When a bi-directional TCP connection is already setup, there is generally no need to set up an additional TCP connection. This is the case both for IPsec and TLS solution. However, if two unidirectional UDP streams are used, both P-CSCF and UE must be able to setup a bi-directional TCP connection.

Deployment

The enhanced IMS access security solution will likely be used in both fixed broadband (TISPAN) and wireless (3GPP) environments. While it may be possible that the same devices could be used to IMS access in both environments, it is believed that PCs will be an important terminal type in fixed broadband environments (in particular short term). Therefore a solution should be chosen that is easily deployed in this environment.

Advantages of TLS

The main advantages of TLS are:

- Proven and mature.This contribution has added two MACs that are used with normal TLS, which is considered a minor update. It should be noted that PSK TLS is not a completely new protocol either, but more an extension of the normal TLS.
- Regular TLS is already available in many SIP client implementations. This makes the deployment of TLS cheap and quick. Furthermore, 3GPP Release 6 UE already supports TLS.
- TLS is easy to integrate to application layer. Software development will be easier and faster.
- No issues when multiple clients used behind a NAT.
- TLS interoperates easily with all existing PC applications, and does not require changes to the operating system.
- Easy and fast deployment of access security into PC environment, which is highly important from business perspective.
- The proposed TLS solutions have no need for PKI.
- Alignment with IETF SIP standard.

A.5 References

[RFC2246] The TLS protocol, IETF, RFC 2246,Proposed standard, updated by RFC3546

[S3-050407] Approaches to TLS based IMS security solutions, Ericsson, TD S3-050407, SA3#39

[S3-050239] Scalability of IMS/TLS server certificate deployment, Ericsson, SA3#38

[33203] Access security for IP based services, 3GPP, TS 33.203v670

[RFC3329] Security Mechanism Agreement for the Session Initiation Protocol (SIP), IETF, RFC 3329, Proposed standard

[PSKTLS] Pre-Shared Key Ciphersuites for Transport Layer Security (TLS), IETF, draft-ietf-tls-psk-09.txt, current RFC Candidate

[RFC2404] Pre- The Use of HMAC-SHA-1-96 within ESP and AH, IETF, RFC 2404, Proposed standard

[S3-050333] NAT device traversal and interoperability issues for IMS Rel-7, T-Mobile, TD S3-050333, SA3#39

[RFC3261] SIP: Session Initiation Protocol, IETF, RFC 3261, Proposed standard

[SIPDRAFT] Managing Client Initiated Connections in the Session Initiation Protocol (SIP), IETF,

draft-ietf-sip-outbound-00

Annex B:
Enabling NAT traversal for signalling messages in the IMS access security framework

Editor's Note: Annex B is based on input document S3-050402 to SA3 #39.

B.1 Introduction

This document proposes a solution that aims at enabling the 3GPP Release 5 and 6 IMS access security mechanisms to operate in scenarios where the UE is located behind a Network Address (and Port) Translator and/or Firewall. This is intended to meet an essential requirement resulting from ETSI TISPAN activities related to its Release 1. A basic feature of the Release 1 architecture is to allow also fixed subscribers to attach to the IMS, including subscribers located

behind a NA(P)T. The goal to provide security enhancements to enable fixed subscribers to attach to an IMS has recently also been approved as a new work item in 3GPP.

In October 2004, BT has already issued a proposal for NAT traversal in the context of IMS access security [S3-040720]. Its basic idea is to use IPSec NAT traversal features (NAT-T), as specified in [RFC3948], to enable the NAT traversal of the Release 5/6 IMS access security solution. Our solution adopts this approach, but discusses it in more detail. However, the discussion is confined to signaling aspects as this is the focus of ETSI TISPAN Release 1. Issues of NAT traversal for media or securing media traffic are out of the scope of our solution and of ETSI TISPAN Release 1.

Furthermore, in this contribution we focus on the issue of traversal of a far-end NAT, i.e. a NAT located at the CPE or access network that is not controlled by the IMS network. Issues of NA(P)T or NA(P)T-PT for address translations between access and core network are not considered.

B.2 Overview

B.2.1 Requirements and Objectives

The design of the solution described in this document was guided by the following requirements/objectives:

- Allow UEs located behind NA(P)Ts to access an IMS based on 3GPP Release 5/6 IMS security concepts.
- It must be possible for multiple UEs behind the same NAT device to access the IMS simultaneously.
- The solution shall modify the existing Release 5/6 IMS access security as specified in TS 33.203 as little as possible.
- The solution shall be based on existing standards as much as possible.
- A mechanism shall be provided that allows both ends, UE and P-CSCF to signal whether they support NAT traversal or not.
- A mechanism shall be provided that allows UE and P-CSCF to find out whether a NAT is located in between UE and P-CSCF or not.
- If no NAT is present between the UE and the P-CSCF, the standard IMS access security procedures shall be applied unmodified.
- The solution shall be compatible with the deployment of a SIP ALG on the P-CSCF as was proposed recently in ETSI TISPAN contribution [06bTD071r1].
- The solution must not introduce any additional security risks compared to the standard IMS access security solution according to [TS 33.203].

B.2.2 Assumptions

- It is assumed that the P-CSCF has a publicly routable IP address

B.2.3 Solution Outline

A basic overview of the initial registration according to 3GPP Release 5/6 IMS access security is given in Error! Reference source not found.. One essential feature of the call flow is that the initial Register message and the following 401 Unauthorized answer stay unprotected (messages 1 and 4 in Error! Reference source not found.), while starting from the second Register Request message on, all messages shall be protected by IPSec (see shaded area comprising messages 5 and 8 and all following messages). The details of the IPSec protection are negotiated using the two messages 1 and 4 (and are confirmed in message 5).

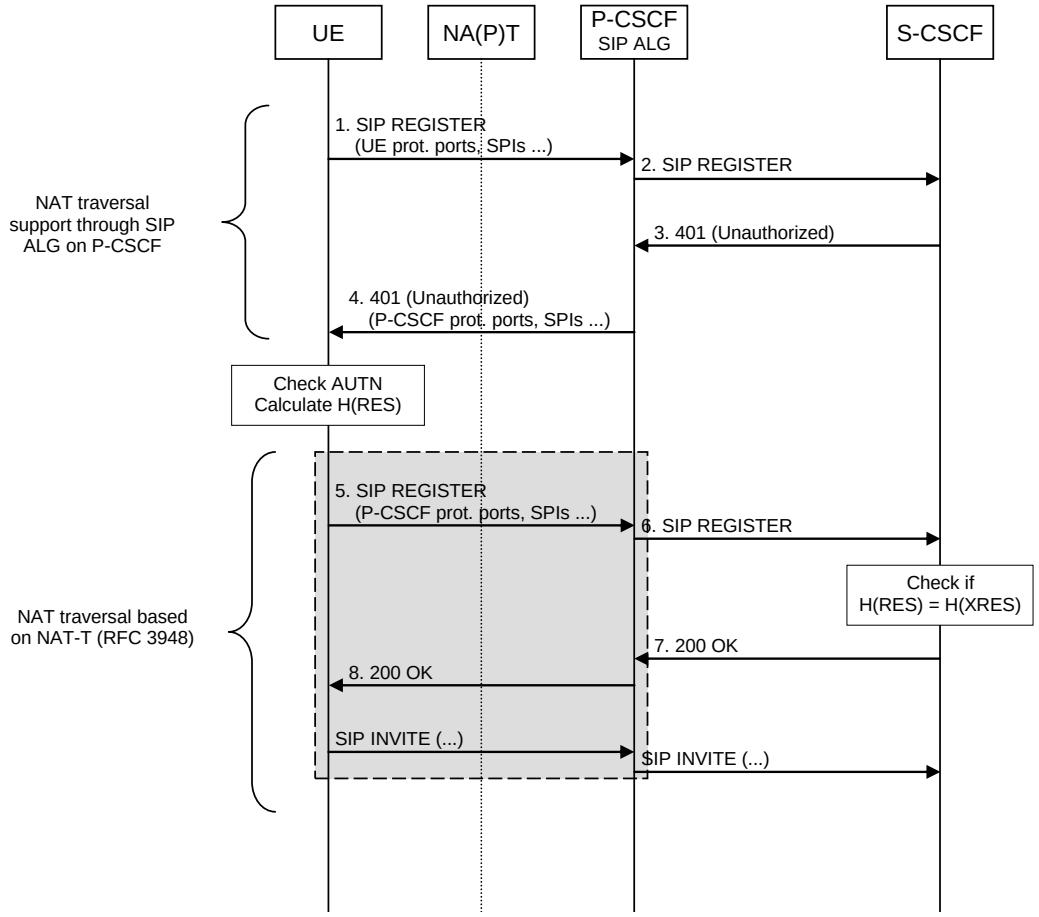


Figure 1: Outline of the IMS Registration procedure

We base our considerations on the existence of a SIP ALG on the P-CSCF as was proposed in ETSI TISPAN contribution [06bTD071r1]. The purpose of this SIP ALG is to perform the necessary modifications in SIP headers and SDP payloads to allow for NAT traversal of signaling and media communication with the UE. With respect to the initial, unprotected SIP messages, we therefore assume that the issue of NAT traversal is handled by the SIP ALG. Later, when the SIP signaling messages are protected by IPSec, UDP encapsulation according to [RFC3948] is used as NAT traversal technique.

Another essential element of the 3GPP Release 5/6 IMS access solution is the fact that two pairs of IPSec SAs are negotiated. These IPSec SAs are bound to IP addresses as well as so-called protected ports which are used to distinguish the different SAs. During an authenticated re-registration, the IPSec SAs are re-negotiated, resulting in a subset of these ports to change. In our solution this mechanism is completely taken as is with no deviation from the standard specification.

Some details of the UDP encapsulation will depend on whether IPSec is used in transport mode (as specified in [TS 33.203]) or tunnel mode. Since our analysis has revealed pros and cons for either mode we will discuss both options in this document.

B.3 Detailed Solution Description

B.3.1 General problems with SIP and NAT (not specific to security)

We assume that the UE is located behind a NAT router that also performs port translation (NAPT), which is quite common in DSL configurations. For simplicity, we will still use the term NAT, denoting both, address and port translation. We further assume that the UE is assigned a private IP address, while the NAT router uses a publicly routable address towards the P-CSCF side.

The problem with SIP signaling (As mentioned above, media traversal is not considered in this annex) and NAT can be summarized as follows (see also [S2-051089]):

(1) When the UE issues a request, the NAT translates the IP source address and the source port and allocates a binding of original and translated address and port. When the response is sent back to the UE, the destination address and port must match the binding in order to be able to pass the NAT. In case of UDP as transport, this will in general not be the case as the UE may send the request from an ephemeral or client port, but the P-CSCF will reply to a well-known or server port.

(2) In addition, any UE-terminating request can only traverse the NAT if it contains a destination address and port that matches an existing NAT binding. Since UDP NAT bindings typically time out quickly in case of signaling inactivity, such a binding must always exist and actively be kept alive.

(3) In the same sense, TCP connections initiated by the P-CSCF will not reach the UE, since the NAT will block TCP connection establishments.

(4) When the UE registers with the S-CSCF it will include its private IP address in the Contact header. Registering a private address does not make sense, since it can not be used to route incoming requests to the UE.

B.3.2 NAT traversal for unprotected messages (not security-specific)

For the initial unprotected Register Request from the UE towards the IMS and the following unprotected 401 "Unauthorized" Response we assume that the SIP ALG deployed in the P-CSCF performs the required procedures. We do not discuss details of the way in which a SIP ALG acts upon the SIP messages, but in general, the SIP ALG will store the public IP address and port information from the UE as received in the IP and UDP/TCP headers as well as the private IP address and port as seen on the SIP message level, like e.g. in the Via and Contact header. It will also typically modifiy the Via and Contact header before forwarding the request, to ensure that the response is routed via the P-CSCF. When the response reaches the P-CSCF it re-writes the SIP headers again and uses the information stored before sending the response towards the UE.

In most configurations, the UE must support symmetric signaling so that the response can traverse the NAT, otherwise no matching binding will be found by the response. Symmetric signaling means that the UE can receive a response on the same port from which the request was sent.

The NAT traversal method for unprotected messages is, as far as we can see, independent of that for protected messages. If this is the case one method could be modified without affecting the other. For example, another standard method to provide NAT traversal for SIP signaling messages is based on the "Symmetric Response Routing" extension specified in [RFC3581].

B.3.3 Detection of NAT traversal capabilities and presence of a NAT (partly security-specific)

Any NAT traversal mechanism shall only be applied in 3G systems if a NAT is really present between UE and P-CSCF. In addition, UEs and P-CSCFs may or may not exhibit NAT traversal capabilities. Therefore it is suggested that both parties signal to each other whether they are able to support NATs in between them and that they detect the presence of a NAT. Signaling the capabilities is preferred, as it allows the P-CSCF to abort an unsuccessful registration already after receiving the first message, without having to signal back to the home network.

The signaling of NAT traversal capabilities can be handled by a header field or header field parameter in the initial SIP request and response message. We propose to enhance the definition of the "mode" parameter of the SIP-Sec-Agree protocol as given in Annex H of [TS 33.203] to accommodate additional values for UDP-encapsulated modes. The modified specification would therefore read as follows:

mode = "mod" EQUAL ("trans" / "tun" / "UDP-enc-trans" / "UDP-enc-tun")

By including appropriate values for the mode parameter, UE and P-CSCF indicate support for the UDP encapsulated NAT traversal. Note that UE and P-CSCF can include multiple mode parameters in the Security-Client, Security-Server or Security-Verify headers.

With respect to the discovery of the presence of a NAT, the P-CSCF can check the source IP address of the received packet against the IP address in the Via header (see also ETSI TISPAN contribution [06bTD070]). If they differ, a NAT is present, and the P-CSCF writes the source IP address of the received packet into the "received" parameter of the Via header. The detection of the presence of a NAT can be performed by the UE by checking the "received" parameter. Note that the "received" parameter is still included in the Via header when the response reaches the UE. That means that the UE can deduce from the presence of a "received" parameter that a NAT is in between the UE and the P-CSCF.

If no NAT is present, none of the NAT specific mechanisms shall be used by either side. If a NAT is present but the UE does not support NAT traversal capabilities, the P-CSCF shall silently discard the request. If a NAT is present and the P-CSCF does not support NAT traversal capabilities, in most cases the UE will not receive a response from the P-CSCF. In case it does (e.g. when TCP was used as transport) and the UE detects that the P-CSCF does not support NAT traversal, the UE shall cancel the registration procedure.

B.3.4 NAT traversal of protected messages (security-specific)

In this section we discuss the NAT traversal of the IPSec protected messages using UDP encapsulation according to [RFC3948]. We only illustrate the message flow, packet contents and essential IPSec SA data in this section in order to point out the underlying mechanism. The important issue of IPSec SA establishment and actual UDP encapsulation handling is discussed in Section 4.

While the current IMS access security standard [TS 33.203] mandates the use of transport mode, we will discuss both, transport and tunnel mode, because each mode has its own advantages and drawbacks as we will also see in Section 4.

B.3.4.1. UDP encapsulation using transport mode

After the first unprotected Register request and reply have been successfully processed, the UE configures two pairs of IPSec SAs and any further messages shall be protected using these SAs. In case of the presence of a NAT and assuming that both, UE and P-CSCF support NAT traversal, UE and P-CSCF switch on the UDP encapsulation mode. The resulting message flow and packet contents are shown in Error! Reference source not found.2. The packet processing at UE and P-CSCF was divided into separate steps in order to show details of the processing steps. Note that this is only a conceptual illustration and does not necessarily represent actual packets in the various processing steps on a machine.

For the message flow and processing of the protected SIP messages described in the following we assume that the SIP ALG does not interfere with the IP addresses and ports in the SIP header fields. The proper routing of the SIP messages is ensured by other means. But the ALG may change other parts of protected SIP messages, e.g. IP addresses and ports in the SDP payloads to enable media routing.

At first the SIP layer at the UE constructs the SIP Register message that it intends to send to the P-CSCF. For proper routing of the response and incoming requests later on, it is important that the UE includes its public IP address in the Via and Contact header of this message. In addition, it must include its protected server port in the Via and Contact field (see considerations below and in Section 3.5). The public IP address can be learned by the UE by evaluating the received parameter contained in the Via header of the (unprotected) "401 Unauthorized" response. The protected server port was selected by the UE at the beginning of the Registration procedure.

When the SIP application layer of the UE hands over the SIP message to the transport layer it indicates the same destination IP address as in the unprotected case. But now the protected ports negotiated before are used for source and destination, instead of the port numbers from the unprotected packets. In the example in Figure 2, UDP is used as transport protocol. This packet is now handed over to IPSec processing which finds appropriate SPD and SAD entries and adds ESP tunnel mode protection to the packet (ESP trailers are not shown in Figure for simplicity). After that, the UDP encapsulation processing adds a UDP header according to [RFC3948]. This includes the use of port 4500 as source and destination ports in the UDP header.

When this packet traverses the NAT, the NAT creates a new binding, which will in most cases be different from the binding used in the initial Registration exchange. In Figure 2, the public source port used by the NAT for the UDP encapsulation header is denoted as port_Uenc.

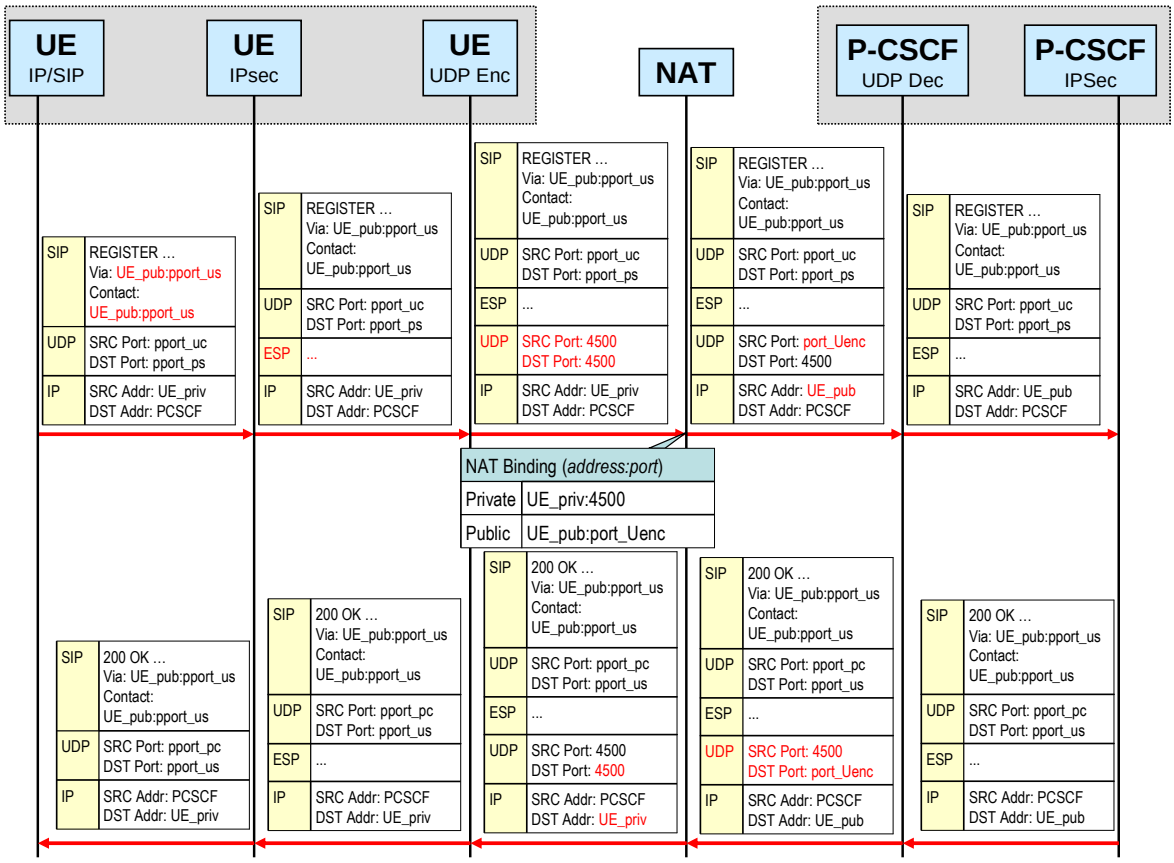


Figure 2: Second Register message

When the packet arrives at the well known port 4500 at the P-CSCF, the P-CSCF performs transport mode decapsulation according to [RFC3948], which means removing the UDP header and adapting some IP header fields. The UDP encapsulation function must also store port port_Uenc and must associate it with the underlying IPSec SA in order to be able to correctly route the response (see Section 4) and all subsequent requests originating from the network.

The normal IPSec processing of the incoming ESP packet follows. It should be noted that the ports and IP addresses found in the ESP protected packet exactly match one of the SAs configured at the P-CSCF. Therefore, IPSec processing proceeds as usual.

Finally, on the SIP level, the P-CSCF will not insert a received parameter, because the UE has used its public IP address in the via header which is the same as the source address in the IP header (it was changed by the NAT). Since there is no discrepancy, no received header will be inserted.

When the P-CSCF eventually sends the response back to the UE it applies normal SIP transport rules, i.e. it inspects the topmost Via header which includes the public IP address of the UE and the protected server port of the UE. This data is handed over to the transport layer. After that the IPSec processing has a matching SA and applies ESP transport protection. The UDP encapsulation that uses the port port_Uenc stored from the incoming message follows next. When this packet arrives at the NAT, a matching binding is available and the NAT translates the packet back to the private address and port used by the UE before. The remaining steps are straightforward and UDP decapsulation and IPSec processing work as expected.

It is important to note that the message flow as described above works equally well with TCP as transport protocol. Since the NAT traversal is completely hidden from the inner transport layer headers, it is immaterial whether UDP or TCP is used. From the point of view of the IPSec processing at both nodes, UE and P-CSCF, the corresponding SAs are selected depending on the transport protocol and whether the message is a request or a response. In this regard, there is no deviation from the standard mechanisms described in [TS 33.203].

In UDP encapsulated transport mode, the IPSec SAs consist of the data as shown in Table 1 and Table 2. Since our focus is on routing issues in the presence of a NAT, we only discuss IP addresses, ports and SPIs. All other IPSec SA data, like algorithms, keys, lifetimes etc. is left out for simplicity. At the P-CSCF (see Table 1) the IP addresses are taken from the source and destination IP addresses as contained in the IP header of the request received. The port

numbers for these SAs are taken from the SIP message received from the UE (for the UE's protected ports) and are selected by the P-CSCF (for the P-CSCF's protected ports). The mode parameter associated with an SA (not shown in Table 1) is set to UDP-Encapsulated-Transport mode, replacing simple Transport mode as used in [TS 33.203].

P-CSCF SA Table				
Selector	SA1	SA2	SA3	SA4
SRC Addr	PCSCF	UE_pub	PCSCF	UE_pub
Dest Addr	UE_pub	PCSCF	UE_pub	PCSCF
SRC Port	pport_pc	pport_uc	pport_ps	pport_us
Dest Port	pport_us	pport_ps	pport_uc	pport_pc
SPI	SPI_us	SPI_ps	SPI_uc	SPI_pc

Table 1: P-CSCF SA Table

At the UE's side (see Table 2), the IMS access security standard does not state anything about the IP address selectors, however, it is assumed that the IP address selectors are also taken from the IP header of the response message, similar to the way in which the P-CSCF behaves. Therefore, the following SA table will result:

UE SA Table				
Selector	SA1	SA2	SA3	SA4
SRC Addr	PCSCF	UE_ priv	PCSCF	UE_ priv
Dest Addr	UE_priv	PCSCF	UE_ priv	PCSCF
SRC Port	pport_pc	pport_uc	pport_ps	pport_us
Dest Port	pport_us	pport_ps	pport_uc	pport_pc
SPI	SPI_us	SPI_ps	SPI_uc	SPI_pc

Table 2: UE SA Table

The SA data is established in full compliance with [TS 33.203], but one can see from the tables that the UE uses its private address in the IP address selector fields, whereas the P-CSCF uses the public address of the UE.

B.3.4.2 UDP encapsulation using tunnel mode

In tunnel mode, the message flow and packet contents are schematically shown in Figure 3Figure . The most salient difference compared to transport mode is an additional inner IP header added right after the ESP header. This implies that both endpoints, UE and P-CSCF now configure two IP addresses, the inner and the outer address. For the P-CSCF we assume that both addresses are the same, namely the public IP address of the P-CSCF. For the UE, the outer address will be the private address, which is typically assigned via DHCP by the local NAT router. As inner address, the UE shall use its public IP address which it learns from the received parameter contained in the response to the first unprotected Register message (see above).

The inner IP address will not be modified by the NAT since it is "hidden" in the ESP tunnel. The outer address is changed by the NAT, so that the P-CSCF will only see the public IP address in the inner and outer header. The handling of the ports and SPIs used for the SAs does not differ compared to the transport mode case. Therefore, the resulting SAs look similar compared to Table 1 and Table 2, except for the additional inner IP address.

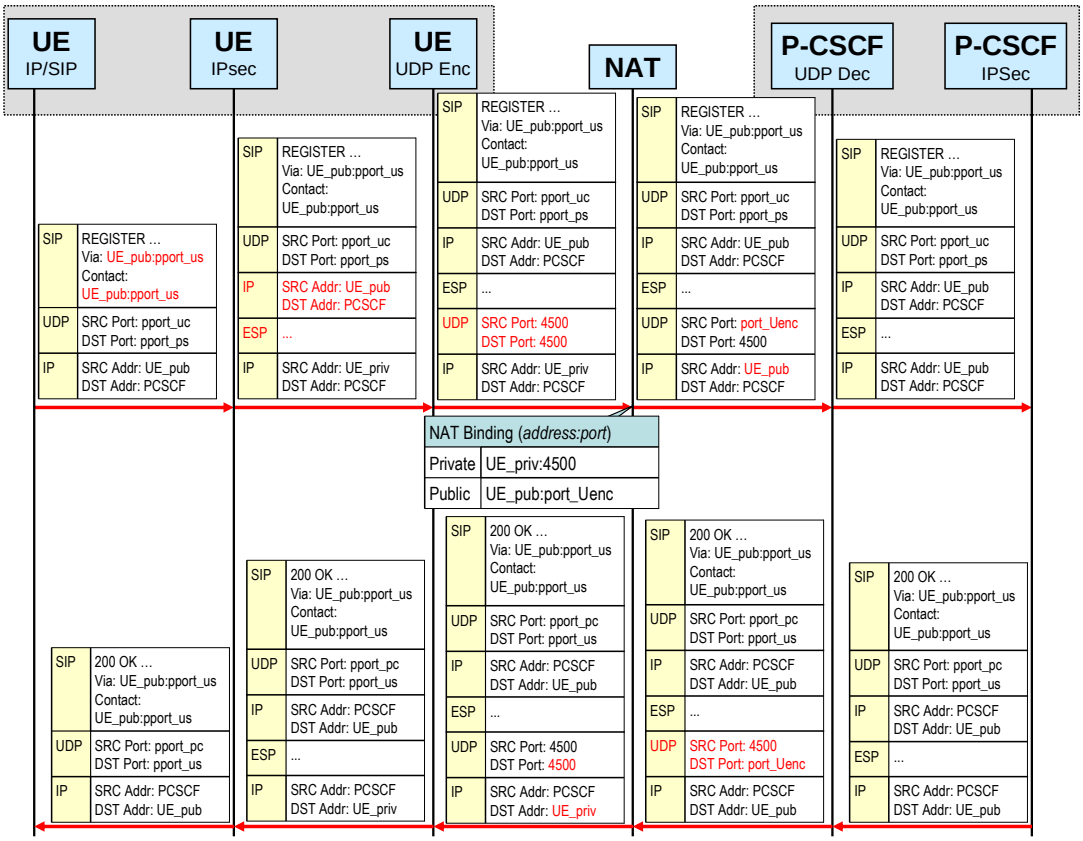


Figure 3: NAT traversal using UDP-encapsulated tunnel mode

B.3.5 Registering a Contact and routing of UE terminating requests (partly security-specific)

In the previous section we have dealt with the routing of protected requests originating from the UE and the corresponding responses. In order for the UE to be able to receive UE terminating requests, the UE must register an appropriate Contact address and port with the S-CSCF. In line with 3GPP specifications, the Contact information given in the first unprotected Register request, which contained the private IP address of the UE as Contact header, is not registered yet. Only when the second protected message yields a successful authentication at the S-CSCF, the Contact header contained therein is registered (see Error! Reference source not found.).

Since it was stated above, that the second Register message shall contain the public IP address and the protected server port of the UE in the Contact header, this data will be registered at the S-CSCF. After this has been performed, an incoming request will make the S-CSCF enter this address and port in the Request URI. The P-CSCF later uses this information to route the incoming request. Since the public IP address and a protected server port is used, the P-CSCF has corresponding SAs established and the normal routing processing including the IPSec handling can proceed. There is no deviation from the standard behaviour.

B.3.6 Keeping the NAT binding alive (not security-specific)

NAT bindings for UDP traffic usually exist only for a short time, typically ranging from 30 seconds up to a few minutes. In order to allow for requests terminating at the UE, the NAT binding must be kept alive during extended periods of inactivity. Since the UDP encapsulation provides such a mechanism it can be reused in this context. According to [RFC3948], a keepalive packet is simply a UDP packet with a single all-ones Byte of payload. Since in our scenario, it is always the UE that is located behind a NAT, only the UE will send keepalive messages. This can be hard-coded into the software and does not have to be negotiated.

B.4 Establishing IPSec SAs and handling of UDP encapsulation

[RFC3948] explicitly states that it is assumed that IKE (either IKEv1 or IKEv2) is used to negotiate UDP encapsulation. It is further stated, that manual configuration is not supported. In fact UDP encapsulation is dynamic in nature, as the port chosen by the NAT and used in the UDP encapsulation header (port_Uenc) can hardly be predicted and must be configured at runtime. In an environment where IKE is used as a means to negotiate UDP encapsulation, this is achieved during IKE phase 1 when the initiator switches to port 4500 (see [RFC3947]). In our case, port_Uenc can only be configured by the time the first protected Register message arrives at the P-CSCF.

Furthermore, one should note that port_Uenc must be considered as part of the SA data of all four SAs established for IMS access security, no matter whether encapsulated transport or tunnel mode is used. This is because the outbound SAs at the P-CSCF (SA1 and SA3 in Table 1) have to know what port to insert as destination port in the UDP encapsulation header. Furthermore, the inbound SAs at the P-CSCF must store port_Uenc in order to determine whether the port used by the NAT has changed (see also discussion below).

While in the presence of IKE, this link between inbound and outbound SAs is provided by IKE itself (IKE "knows" what pair(s) of SAs it negotiates and has a means to store this relationship in the SAD), in our case the only entity that knows that the four SAs are related and that is capable of configuring port_Uenc, is the SIP application at the P-CSCF. Consequently, the SIP application at the P-CSCF (or a separate application with an appropriate interface to the SIP application) must somehow receive the information of port_Uenc and configure it into the IPSec SAs. It is important to note that port_Uenc only has to be configured dynamically at the P-CSCF's side. The UE is not affected by any NAT translation of the UDP encapsulating port. It will always see port 4500 for both, source and destination ports.

Another issue to consider is the fact that according to [RFC3947] and [RFC3948], UDP encapsulated packets for ESP and IKE must use the same well know port 4500. They are distinguished by a payload starting with either four zero octects (IKE) or a different value (the ESP SPI). Thus, in standard scenarios, an IKE NAT-T capable daemon listens on port 4500 and demultiplexes IKE and ESP traffic. In this configuration there may be implications for the implementation in case IKE is also used on the same network interface of a P-CSCF, since the standard assumes the same port number for UDP encapsulated IKE and ESP traffic.

Finally, there are subtle differences between tunnel and transport mode with respect to checksum calculations, which may also influence design decisions. In tunnel mode, the UDP/TCP checksum, which includes the IP addresses of the tunneled IP header, are not affected by the NAT, since the NAT does not change the inner IP address. In transport mode, the IP addresses that are used for the checksum calculation are changed by the NAT, so that the checksum will not be successfully verified.

Following these considerations we present two UDP encapsulation based approaches to the NAT traversal problem which are described in the following subsections. The first proposes not to use built in IPSec features for UDP encapsulation processing but to use a separate application, called the UDP encapsulation function. This application is either integrated into the SIP application at the P-CSCF or consists of a separate application that has a communication link to the SIP application. The second approach uses the UDP encapsulation features of IPSec and assumes that the IPSec processing and the SAD-interface is capable of providing all required hooks to the SIP application in order to properly configure the SA and UDP encapsulation related data. For reasons described below, the first approach uses IPSec tunnel mode, while the second approach uses transport mode.

B.4.1 Using a separate UDP encapsulation function and UDP encapsulated tunnel mode

An outline of the solution approach is illustrated in Figure 4. We show a separate function that handles UDP encapsulation on the P-CSCF. The main advantage of this approach is that it allows a modular add-on of the encapsulation functionality to the IMS Release 5 solution.

After the UE has sent the first unprotected Register message, and the P-CSCF has received the response from the S-CSCF, the P-CSCF configures two pairs of IPSec SAs at the IPSec layer as in [TS 33.203] but this time using IPSec tunnel mode. In addition, the P-CSCF also informs the UDP encapsulation function about the IP addresses and SPIs used for each SA established. This results in an UDP encapsulation table as shown in Figure 4.

The UDP encapsulation table contains for each SA the source and destination IP addresses, the source and destination ports as contained in the UDP encapsulating header and the SPI used. At this stage of the protocol execution, the table is

still incomplete, since port_Uenc is not known yet. Assuming that the UE sends its UDP encapsulated packets to the well-known port 4500 and the UDP encapsulation function listens on that port, the first protected Register message from the UE will contain the port_Uenc as source port in the UDP header (message 5 in Figure 4). The UDP encapsulation function can now identify the SA used by means of the SPI and destination address, which is supposed to be unique by definition. It takes port_Uenc and configures it in the UDP encapsulation table at the appropriate places, i.e. at all related SAs (see Figure 4). Note that the SPI can always be read from the ESP header, even if encryption is applied.

The essential idea of this approach is now that the UDP encapsulation function uses the information from the UDP encapsulation table to perform the UDP encapsulation for NAT traversal. For example, for the 200 OK response (message 6 in Figure 4), assuming that UDP is used as transport protocol for SIP, SA3 will be used. Thus, taking the destination address of the packet and the SPI will together yield a unique entry in the UDP encapsulation table enabling the UDP encapsulation function to add the appropriate destination port port_Uenc.

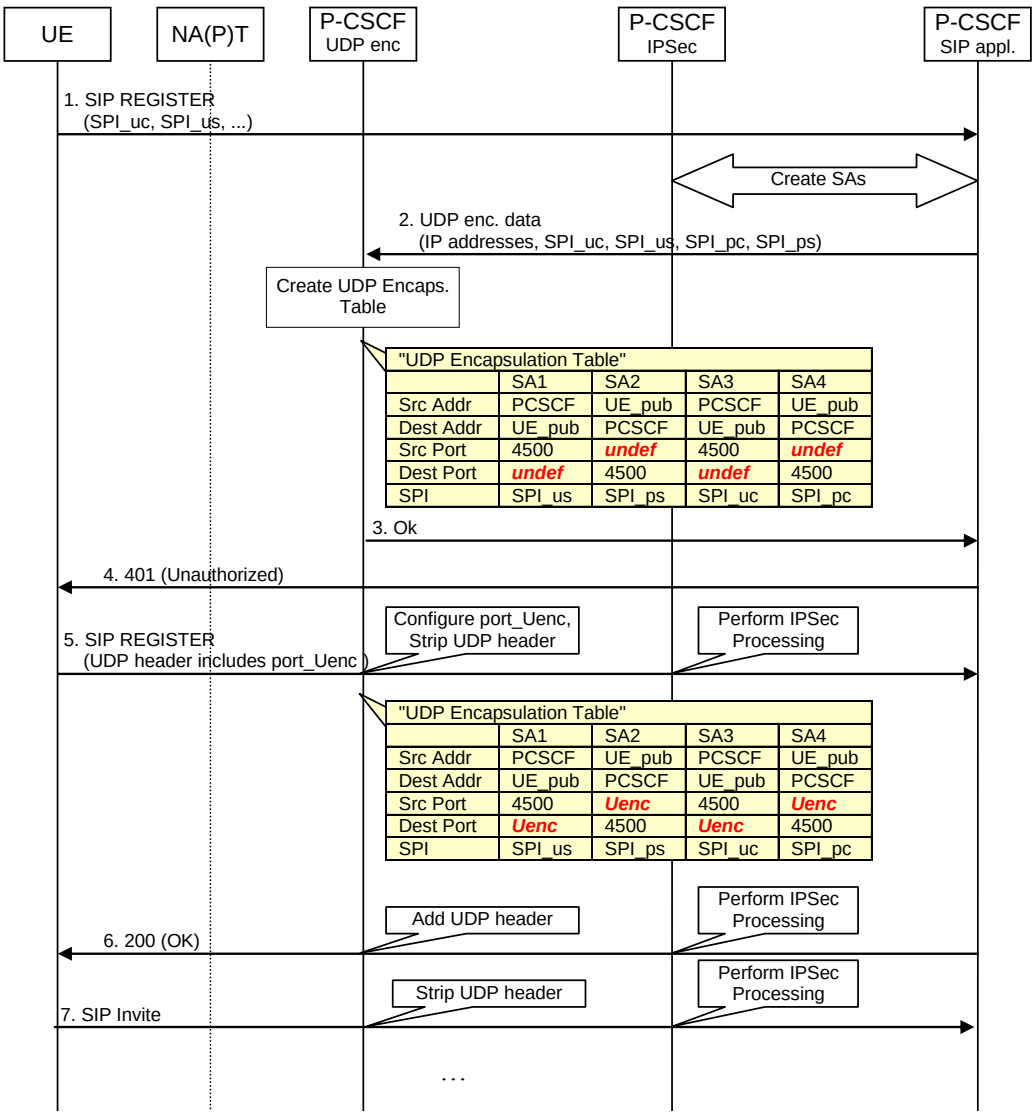


Figure 4: Using a separate UDP encapsulating application and IPSec tunnel mode

For inbound processing, when UDP encapsulated packets are received (e.g. message 7 in Figure 4), the UDP encapsulation function checks whether a matching table entry exists. If yes, it just strips the UDP header and forwards the packet to IPSec processing. In case the packet was a bogus packet created by an attacker using valid combinations of IP addresses, ports and SPI, the following IPSec processing will fail and drop the packet. In this regard there is no difference to the case without UDP encapsulation.

If the UDP encapsulation table does not have a matching binding, the UDP encapsulation function must drop the packet. It should be noted, that the NAT-T standard ([RFC3947] and [RFC3948]) mandates that IP address and port selectors shall be adapted in case of a NAT changing its binding, e.g. due to re-boot. However, this requires that the

IPSec processing was executed successfully. In our case, the UDP encapsulation function cannot check whether the IPSec processing will be successful after forwarding a packet with modified source port and address to it. Thus, the case of changing NAT bindings must be excluded. In practice this is not considered to be too strong a constraint, as the case of a re-booting NAT can be seen as a very rare event.

As a prerequisite for the mechanism to work, the combination of (SPI, destination IP address) for messages towards a UE must be unique. But in general, SPI clashes at two different UEs cannot be prevented. If these UEs are located behind the same NAT, and thus are assigned the same public IP address, the combination (SPI, destination IP address) is not unique. Consequently, the P-CSCF, when receiving an initial Register request with a combination of (SPI, destination IP address) that is already used for an SA at the P-CSCF's side, must reject the registration attempt and prompt the UE to choose a new SPI (see also discussion in Section 5).

Another important advantage of the selection of tunnel mode instead of transport mode is that the verification of the UDP/TCP checksum does not create any problems as it is completely included and protected inside the ESP tunnel.

It should be noted that the IPSec application at the P-CSCF does not apply any UDP encapsulation features, rather it operates in a standard mode without the extensions described in [RFC3948]. On the other hand, we assume that the UDP encapsulation function at the UE's side uses the IPSec UDP encapsulation feature. Therefore, since the UDP encapsulation at the UE will automatically send the UDP encapsulated packets to port 4500 at the P-CSCF, the UDP encapsulation function on the P-CSCF must listen on port 4500 and no IKE daemon must run on that interface on the P-CSCF.

B.4.2 Using IPSec built-in UDP encapsulation features and UDP encapsulated transport mode

In the approach discussed in this subsection we assume an IPSec implementation with integrated UDP encapsulation functionality. Due to the fact, that the checksum correction will then be performed by the IPSec implementation, as mandated in [RFC3948], transport mode can be used instead of tunnel mode. Thus, one of the main advantages of this approach is its relative efficiency compared to tunnel mode. However, it requires an IPSec implementation that provides the UDP encapsulation functionality and the possiblitity to integrate such functionality into the IMS framework. The resulting high level call flow is shown in Figure 5.

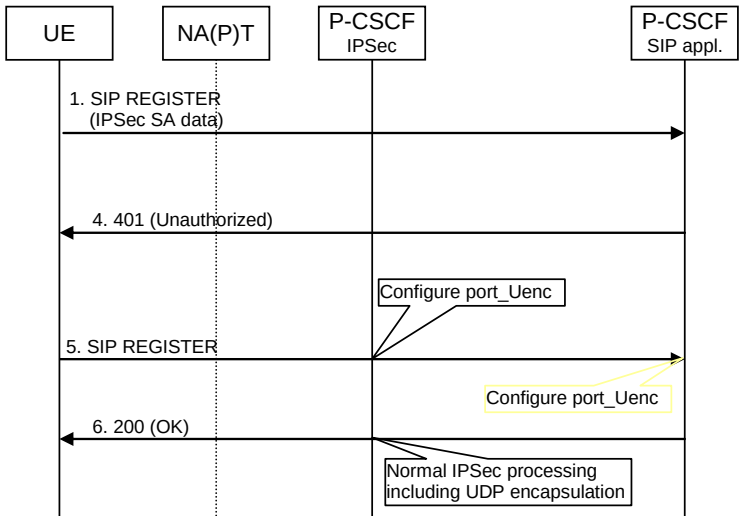


Figure 5: Using IPSec built-in UDP encapsulation features and UDP encapsulated transport mode

One of the important points to consider is the question how the UDP encapsulation port "port_Uenc" can be configured into the SAs at the P-CSCF. This port is only known when the first protected message arrives at the P-CSCF. Depending on the implementation of IPsec with integrated UDP encapsulation, when the SAs are created in the SA database by the P-CSCF application the latter may also add information that the four SAs relating to one registration belong together, and it may be possible to provide the port "port_Uenc" to all related SAs when the protected REGISTER message arrives, without again involving the P-CSCF application. Alternatively, the P-CSCF application may dynamically enter the port "port_Uenc" to all related SAs when the protected REGISTER message arrives.

B.5 Multiple UEs behind the same NAT

B.5.1 Implications from the use of a common (public) IP address for multiple UEs

Multiple UEs behind the same NAT is a common scenario in DSL configurations (see 6) and the solution must be able to cope with it. Typically, such a situation implies that the NAT uses the same public IP address for both UEs. In addition, it can not be avoided that the UEs select the same port number for either one or both of the protected ports. In this case, the P-CSCF must ensure that unambiguous Security Associations are established with respect to the IP addresses and ports as selectors.

[TS 33.203] already excludes that a Registration is accepted by the P-CSCF if the pair (UE_IP_address, UE_protected_client_port) included in the Register message is used in an SA in the SA table at the P-CSCF. Such a registration attempt must be answered by the P-CSCF with an appropriate error message. Consequently, the case where the two UEs behind the same NAT use the same protected client port is already covered by [TS 33.203].

In addition, it must be ensured that no clash occurs in case the two UEs behind the same NAT select the same protected server port. There seem to be two options to address this:

(1) The P-CSCF rejects the attempt to register using an IP address and protected UE server port that is already used in an SA in the SA table. This is similar to the case of a clash with the protected client port.

(2) Alternatively, the P-CSCF simply selects at its side a protected client port that is different from the one used in the already existing SA. This will make the selector values in the new SA unambiguous.

Case (2) seems to be the simplest option, since it does not require an error message and additional round trip. On the other hand, in option (2) two UEs register a Contact with the same IP address and protected server port. While this does not seem to be a problem from a theoretical point of view in the context considered here – the correct routing of messages to the UEs is ensured by the UDP encapsulation using different ports – it is for further study whether there are implications elsewhere.

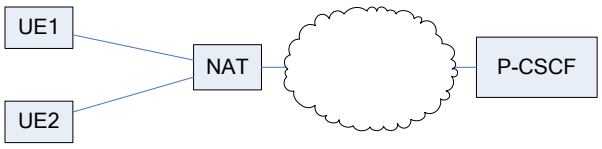


Figure 6: Multiple UEs behind the same NAT

B.6 References

[06bTD071r1] Ericsson, France Telecom, Contribution to ETSI TISPAN WG2: "NA(P)T-PT and Hosted NAT procedures (Stage 2)", 06bTD071r1, June 2005.

[06bTD070] Ericsson, France Telecom, Contribution to ETSI TISPAN WG3: "Hosted NAT detection", 06bTD070, June 2005.

[S2-051089] Nokia, Contribution to 3GPP SA2#46: "NAT traversal for IMS", S2-051089, May 2005.

[S3-040720] BT Group, Contribution to SA3#35: "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", S3-040720, October 2004. See also related Change Request S3-040721.

[RFC3261] J. Rosenberg et al.: "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[RFC3329] J. Arkko et al.: "Security Mechanism Agreement for the Session Initiation Protocol (SIP)", RFC 3329, January 2003.

[RFC3581] J. Rosenberg, H. Schulzrinne: "An Extension to the Session Initiation Protocol (SIP) for Symmetric Response Routing", RFC 3581, August 2003.

[RFC3947] T. Kivinen et al.: "Negotiation of NAT-Traversal in the IKE", RFC 3947, January 2005.

[RFC3948] A. Huttunen et al.: "UDP Encapsulation of IPsec ESP Packets", RFC 3948, January 2005.

Annex C:
Generic Network Tunnel (GNT) for NGN

The Generic Network Tunnel solution is described in ETSI TISPAN document 06bTD137.

Annex D:
Enabling NAT traversal for signaling messages in the IMS access security framework

Editor's Note: Annex D is based on input document S3-050533 to SA3 #40, and shows proposed changes to 33.203 for enabling NAT Traversal based on UDP encapsulated IPsec.

***** BEGIN SET OF CHANGES *****

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TS 33.102: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security Architecture".

[2] 3GPP TS 22.228: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Service Requirements for the IP Multimedia Core Network".

[3] 3GPP TS 23.228: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia (IM) Subsystem".

[4] 3GPP TS 21.133: "3rd Generation Partnership Project; T Technical Specification Group Services and System Aspects; Security Threats and Requirements ".

[5] 3GPP TS 33.210: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Network domain security; IP network layer security".

[6] IETF RFC 3261 "SIP: Session Initiation Protocol".

[7] 3GPP TS 21.905: "3rd Generation Partnership Project: Technical Specification Group Services and System Aspects; Vocabulary for 3GPP specifications".

[8] 3GPP TS 24.229: "3rd Generation Partnership Project: Technical Specification Group Core Network; IP Multimedia Call Control Protocol based on SIP and SDP".

[9] 3GPP TS 23.002: "3rd Generation Partnership Project: Technical Specification Group Services and System Aspects, Network Architecture".

[10] 3GPP TS 23.060: "3rd Generation Partnership Project: Technical Specification Group Services and System Aspects, General Packet Radio Service (GPRS); Service Description".

[11] 3GPP TS 24.228: "3rd Generation Partnership Project: Technical Specification Group Core Network; Signalling flows for the IP multimedia call control based on SIP and SDP".

[12] IETF RFC 2617 (1999) "HTTP Authentication: Basic and Digest Access Authentication".

[13] IETF RFC 2406 (1998) "IP Encapsulating Security Payload (ESP)".

[14] IETF RFC 2401 (1998) "Security Architecture for the Internet Protocol".

[15] IETF RFC 2403 (1998) "The Use of HMAC-MD5-96 within ESP and AH".

[16] IETF RFC 2404 (1998) "The Use of HMAC-SHA-1-96 within ESP and AH".

[17] IETF RFC 3310 (2002): "HTTP Digest Authentication Using AKA". April, 2002.

[18] IETF RFC 3041 (2001): "Privacy Extensions for Stateless Address Autoconfiguration in IPv6".

[19] IETF RFC 2402 (1998): "IP Authentication Header".

[20] IETF RFC 2451 (1998): "The ESP CBC-Mode Cipher Algorithms ".

[21] IETF RFC 3329 (2002): "Security Mechanism Agreement for the Session Initiation Protocol (SIP)".

[22] IETF RFC 3602 (2003): " The AES-CBC Cipher Algorithm and Its Use with IPsec".

[23] IETF RFC 3263 (2002): "Session Initiation Protocol (SIP): Locating SIP Servers".

[24] 3GPP TS 33.310: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Network Domain Security (NDS); Authentication Framework (AF)".

[25] 3GPP TR 33.978: "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security Aspects Of Early IMS".

[26] IETF RFC 3947 (2005): "Negotiation of NAT-Traversal in the IKE".

[27] IETF RFC 3948 (2005): "UDP Encapsulation of IPsec ESP Packets".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply.

Authenticated (re-) registration: A registration i.e. a SIP register is sent towards the Home Network which will trigger a authentication of the IMS subscriber i.e. a challenge is generated and sent to the UE.

Confidentiality: The property that information is not made available or disclosed to unauthorised individuals, entities or processes.

Data integrity: The property that data has not been altered in an unauthorised manner.

Data origin authentication: The corroboration that the source of data received is as claimed.

Entity authentication: The provision of assurance of the claimed identity of an entity.

Key freshness: A key is fresh if it can be guaranteed to be new, as opposed to an old key being reused through actions of either an adversary or authorised party.

ISIM – IM Subscriber Identity Module: For the purposes of this document the ISIM is a term that indicates the collection of IMS security data and functions on a UICC. The ISIM may be a distinct application on the UICC.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply, TS 21.905 [7] contains additional applicable abbreviations:

AAA	Authentication Authorisation Accounting
AKA	Authentication and key agreement
CSCF	Call Session Control Function
HSS	Home Subscriber Server
IM	IP Multimedia
IMPI	IM Private Identity
IMPU	IM Public Identity
IMS	IP Multimedia Core Network Subsystem
ISIM	IM Services Identity Module
MAC	Message Authentication Code
ME	Mobile Equipment
NAPT	Network Address and Port Translation
NAT	Network Address Translation
SA	Security Association
SEG	Security Gateway
SDP	Session Description Protocol
SIP	Session Initiation Protocol
UA	User Agent

***** END SET OF CHANGES *****

***** BEGIN SET OF CHANGES *****

Annex A: Enhancements to the access security for IP based services to enable NAT traversal for signaling messages

Note: section A.x (x= 1, 2, …) in this annex corresponds to section x in the body of this specification.

Editor's note: although this annex is new and therefore the entire text should be marked as revision, the text below shows revision marks only when it differs from the corresponding text in the body of this specification. This is meant to help the reader to better understand the differences between the text in this annex and the specification in the body.

A.1 Scope

It is assumed for the purposes of this annex that a NAT device may be located between the UE and the P-CSCF. Only NATs outside the borders of an IMS network are considered, i.e. NATs are assumed to be located at the subscriber's site or in the access network. If there are multiple NATs in either of these locations, it is assumed that their effect sums up in such a way that they can be treated as a single NAT so that the mechanisms described below are still valid.

In this annex enhancements to sections 4 through 8 of this specification are specified that allow a UE and a P-CSCF to detect whether they are located behind a NAT device, to inform each other about their NAT traversal capabilities, and, if there is a NAT present, to securely communicate. If there is no NAT device present, the procedures of sections 6, 7 and 8 apply. Examples of subscribers who are, in general, located behind a NAT device include subscribers accessing IMS via a DSL line.

Furthermore, this specification is restricted to the treatment of NAT traversal for signalling messages. Measures required for NAT traversal of media data is not considered.

It should be noted that many NAT routers in residential sites do also apply port translation, which is typically denoted as Network Address and Port Translation (NAPT). For reasons of simplicity the term NAT is used, no matter whether only address or address and port translation is actually applied.

A.2 References

Additional references used in this section were incorporated directly into section 2.

A.3 Definitions, symbols and abbreviations

Additional definitions, symbols and abbreviatios nused in this section were incorporated directly into section 3. s

A.4 Overview of the security architecture

The text in section 4 applies without changes.

A.5 Security features

The text in section 5 applies without changes.

A.6 Security mechanisms

A.6.1 Authentication and key agreement

The text in section 6.1 applies without changes.

A.6.2 Confidentiality mechanisms

If the local policy in P-CSCF requires the use of IMS specific confidentiality protection mechanism between UE and P-CSCF, IPsec ESP as specified in RFC 2406 [13] shall provide confidentiality protection of SIP signalling between the UE and the P-CSCF, protecting all SIP signalling messages at the IP level. IPSec ESP general concepts on Security Policy management, Security Associations and IP traffic processing as described in reference RFC 2401 [14] shall also be considered. ESP confidentiality shall be applied ~~in transport mode~~ between UE and P-CSCF either in transport mode if no NAT is present, or – if NAT traversal shall be supported – in UDP encapsulated tunnel mode.

The method to set up ESP security associations (SAs) during the SIP registration procedure is specified in clause A.7. As a result of an authenticated registration procedure, two pairs of unidirectional SAs between the UE and the P-CSCF all shared by TCP and UDP, shall be established in the P-CSCF and later in the UE. One SA pair is for traffic between a client port at the UE and a server port at the P-CSCF and the other SA is for traffic between a client port at the P-CSCF and a server port at the UE. For a detailed description of the establishment of these security associations see clause A.7.

The encryption key CK_{ESP} is the same for the two pairs of simultaneously established SAs. The encryption key CK_{ESP} is obtained from the key CK_{IM} established as a result of the AKA procedure, specified in clause A.6.1, using a suitable key expansion function.

The encryption key expansion on the user side is done in the UE. The encryption key expansion on the network side is done in the P-CSCF.

A.6.3 Integrity mechanisms

IPsec ESP as specified in reference RFC 2406 [13] shall provide integrity protection of SIP signalling between the UE and the P-CSCF, protecting all SIP signalling messages at the IP level. IPSec ESP general concepts on Security Policy management, Security Associations and IP traffic processing as described in reference RFC 2401 [14] shall also be considered. ESP integrity shall be applied ~~in transport mode~~ between UE and P-CSCF either in transport mode if no NAT is present or – if NAT traversal shall be supported – in UDP encapsulated tunnel mode.

The method to set up ESP security associations (SAs) during the SIP registration procedure is specified in clause A.7. As a result of an authenticated registration procedure, two pairs of unidirectional SAs between the UE and the P-CSCF, all shared by TCP and UDP, shall be established in the P-CSCF and later in the UE. One SA pair is for traffic between a client port at the UE and a server port at the P-CSCF and the other SA is for traffic between a client port at the P-CSCF and a server port at the UE. For a detailed description of the establishment of these security associations see clause A.7.

The integrity key IK_{ESP} is the same for the two pairs of simultaneously established SAs. The integrity key IK_{ESP} is obtained from the key IK_{IM} established as a result of the AKA procedure, specified in clause A.6.1, using a suitable key expansion function. This key expansion function depends on the ESP integrity algorithm and is specified in Annex I of this specification.

The integrity key expansion on the user side is done in the UE. The integrity key expansion on the network side is done in the P-CSCF.

The anti-replay service shall be enabled in the UE and the P-CSCF on all established SAs.

A.6.4 Hiding mechanisms

The text in section 6.4 applies without changes.

A.6.5 CSCF interoperating with proxy located in a non-IMS network

The text in section 6.5 applies without changes.

A.7 Security association set-up procedure

The security association set-up procedure is necessary in order to decide what security services to apply and when the security services start. In the IMS authentication of users is performed during registration as specified in clause A.6.1. Subsequent signalling communications in this session will be integrity protected based on the keys derived during the authentication process.

A.7.1 Security association parameters

For protecting IMS signalling between the UE and the P-CSCF it is necessary to agree on shared keys that are provided by IMS AKA, and a set of parameters specific to a protection method. The security mode setup (cf. clause A.7.2) is used to negotiate the SA parameters required for IPsec ESP with authentication and confidentiality, in accordance with the provisions in clauses A.5.1.3 and A.6.2.

The SA parameters that shall be negotiated between UE and P-CSCF in the security mode set-up procedure are:

- **Encryption algorithm**

  The encryption algorithm is either DES-EDE3-CBC as specified in RFC 2451 [20] or AES-CBC as specified in RFC 3602 [22] with 128 bit key.

  Both encryption algorithms shall be supported by both, the UE and the P-CSCF.

- **Integrity algorithm**

NOTE: What is called "authentication algorithm" in RFC 2406 [13] is called "integrity algorithm" in this specification in order to be in line with the terminology used in other 3GPP specifications and, in particular, to avoid confusion with the authentication algorithms used in the AKA protocol.

  The integrity algorithm is either HMAC-MD5-96 [15] or HMAC-SHA-1-96 [16].

  Both integrity algorithms shall be supported by both, the UE and the P-CSCF as mandated by RFC 2406 [13]. In the unlikely event that one of the integrity algorithms is compromised during the lifetime of this specification, this algorithm shall no longer be supported.

NOTE: If only one of the two integrity algorithms is compromised then it suffices for the IMS to remain secure that the algorithm is no longer supported by any P-CSCF. The security mode set-up procedure (cf. clause 7.2) will then ensure that the other integrity algorithm is selected.

- **Mode**

  The IPSec SA mode of operation shall depend on whether the UE is located behind a NAT device or not. If the UE is located behind a NAT device UDP encapsulated tunnel mode according to [26] shall be used. Otherwise transport mode shall be used. The security mode setup (cf. clause A.7.2) allows the P-CSCF to detect whether the UE is located behind a NAT or not.

- **SPI (Security Parameter Index)**

  The SPI is allocated locally for inbound SAs. The triple (SPI, destination IP address, security protocol) uniquely identifies an SA at the IP layer. The UE shall select the SPIs uniquely, and different from any SPIs that might be used in any existing SAs (i.e. inbound and outbound SAs). The SPIs selected by the P-CSCF shall be different than the SPIs sent by the UE, cf. clause 7.2. In an authenticated registration, the UE and the P-CSCF each select two SPIs, not yet associated with existing inbound SAs, for the new inbound security associations at the UE and the P-CSCF respectively.

NOTE: This allocation of SPIs ensures that protected messages in the uplink always differ from protected messages in the downlink in, at least, the SPI field. This thwarts reflection attacks. When several applications use IPsec on the same physical interface the SIP application should be allocated a separate range of SPIs.

The following SA parameters are not negotiated:

- Life type: the life type is always seconds;
- SA duration: the SA duration has a fixed length of 2^{32}-1;

NOTE: The SA duration is a network layer concept. From a practical point of view, the value chosen for "SA duration" does not impose any limit on the lifetime of an SA at the network layer. The SA lifetime is controlled by the SIP application as specified in clause A.7.4.

~~- Mode: transport mode;~~

- Key length: the length of the integrity key IK_{ESP} depends on the integrity algorithm. It is 128 bits for HMAC-MD5-96 and 160 bits for HMAC-SHA-1-96.
- Key length: the length of the encryption key depends on the encryption algorithm. The entropy of the key shall at least be 128 bits.

Selectors if no NAT is present:

Cf. section 7.1

Selectors if a NAT is present:

The security associations (SA) have to be bound to specific parameters (selectors) of the SIP flows between UE and P-CSCF, i.e. source and destination IP addresses, transport protocols that share the SA, and source and destination ports.

~~- IP addresses are bound~~ If a NAT is present, it is assumed that the UE is configured locally with a (e.g. private) IP address. When the UE communicates with the P-CSCF via the NAT device, the NAT allocates a binding, mapping the local IP address to ~~two pairs of SAs, as~~ a publicly routable IP address (called public IP address in the sequel) and perhaps also mapping the source port used in ~~clause 6.3, as follows:~~ the UDP or TCP packet to another port number.

- IP addresses:
  - inbound SA at the P-CSCF:
    The source and destination IP addresses associated with the SA are identical to those in the header of the IP packet in which the initial SIP REGISTER message was received by the P-CSCF.
  - outbound SA at the P-CSCF:
    ~~the~~ The source IP address bound to the outbound SA equals the destination IP address bound to the inbound SA;
    the destination IP address bound to the outbound SA equals the source IP address bound to the inbound SA.

NOTE: This implies that the source and destination IP addresses in the header of the IP packet in which the protected SIP REGISTER message was received by the P-CSCF need to be the same as those in the header of the IP packet in which the initial SIP REGISTER message was received by the P-CSCF.

NOTE: This further implies that the source address in the inbound SA and the destination address in the outbound SA at the P-CSCF equals the public IP address of the UE.

- outbound SA at the UE:
  The source IP address bound to the outbound SA equals the public IP address of the UE. The public IP address is learned by the UE from the received parameter in the Via header in the 401 Unauthorized response to the initial unprotected REGISTER Request (cf Section A.7.2).
  The destination IP address bound to the outbound SA equals the destination IP address in the header of the IP packet in which the initial SIP REGISTER was sent to the P-CSCF.
- inbound SA at the UE:
  The source IP address bound to the inbound SA equals the destination IP address bound to the outbound SA; the destination IP address bound to the inbound SA equals the source IP address bound to the outbound SA.

NOTE: For the handling of the outer IP header in UDP encapsulated tunnel mode, see section on "Data related to the use of UDP encapsulated tunnel mode" below.

- The transport protocol selector shall allow UDP and TCP.

- Ports:

1. The P-CSCF associates two ports, called *port_ps* and *port_pc*, with each pair of security assocations established in an authenticated registration. The ports *port_ps* and *port_pc* are different from the standard SIP ports 5060 and 5061. No unprotected messages shall be sent from or received on the ports *port_ps* and *port_pc*. From a security point of view, unprotected messages may be received on any port which is different from the ports *port_ps* and *port_pc*. The number of the ports *port_ps* and *port_pc* are communicated to the UE during the security mode set-up procedure, cf. clause 7.2. These ports are used with both, UDP and TCP. The use of these ports may differ for TCP and UDP, as follows:

   UDP case: the P-CSCF receives requests and responses protected with ESP from any UE on the port *port_ps* (the"protected server port"). The P-CSCF sends requests and responses protected with ESP to a UE on the port *port_pc* (the "protected client port").

   TCP case: the P-CSCF, if it does not have a TCP connection towards the UE yet, shall set up a TCP connection from its *port_pc* to the port *port_us* of the UE before sending a request to it..

NOTE: Both the UE and the P-CSCF may set up a TCP connection from their client port to the other end's server port on demand. An already existing TCP connection may be reused by both the P-CSCF or the UE; but it is not mandatory.

NOTE: The protected server port *port_ps* stays fixed for a UE until all IMPUs from this UE are de-registered. It may be fixed for a particular P-CSCF over all UEs, but there is no need to fix the same protected server port for different P-CSCFs.

NOTE: The distinction between the UDP and the TCP case reflects the different behaviour of SIP over UDP and TCP, as specified in section 18 of RFC 3261 [6].

NOTE: The handling of the protected ports is the same, irrespective of whether transport or UDP encapsulated tunnel mode is used.

2. The UE associates two ports, called *port_us* and *port_uc*, with each pair of security assocations established in an authenticated registration. The ports *port_us* and *port_uc* are different from the standard SIP ports 5060 and 5061. No unprotected messages shall be sent from or received on the ports *port_us* and *port_uc*. From a security point of view, unprotected messages may be received on any port which is different from the ports *port_us* and *port_uc*. The number of the ports *port_us* and *port_uc* are communicated to the P-CSCF during the security mode set-up procedure, cf. clause 7.2. These ports are used with both, UDP and TCP. The use of these ports may differ for TCP and UDP, as follows:

   UDP case: the UE receives requests and responses protected with ESP on the port *port_us* (the"protected server port"). The UE sends requests and responses protected with ESP on the port *port_uc* (the "protected client port").

   TCP case: the UE, if it does not have a TCP connection towards the P-CSCF yet, shall set up a TCP connection to the port *port_ps* of the P-CSCF before sending a request to it.

NOTE: Both the UE and the P-CSCF may set up a TCP connection from their client port to the other end's server port on demand. An already existing TCP connection may be reused by both the P-CSCF or the UE, but it is not mandatory.

NOTE: The protected server port *port_us* stays fixed for a UE until all IMPUs from this UE are de-registered.

NOTE: The distinction between the UDP and the TCP case reflects the different behaviour of SIP over UDP and TCP, as specified in section 18 of RFC 3261 [6]

NOTE: The handling of the protected ports is the same, irrespective of whether transport or UDP encapsulated tunnel mode is used.

3. The P-CSCF is allowed to receive only REGISTER messages and error messages on unprotected ports. All other messages not arriving on a protected port shall be either discarded or rejected by the P-CSCF.

4. The UE is allowed to receive only the following messages on an unprotected port:

   - responses to unprotected REGISTER messages;
   - error messages.

   All other messages not arriving on a protected port shall be rejected or silently discarded by the UE.

Data related to the use of UDP encapsulated tunnel mode

- Tunnel endpoint addresses and header construction for tunnel mode:

  In case UDP encapsulated tunnel mode is selected, an "outer" IP header is added to protected packets exchanged between UE and P-CSCF, following the rules of tunnel mode processing according to [14]. While the IP addresses of the inner IP header are as specified above in the section about "Selectors", the IP addresses of the outer IP header shall be selected as follows:

  - P-CSCF:
  For the outbound SA at the P-CSCF the source address shall be the IP address of the P-CSCF, the destination address shall be the public IP address of the UE. For the inbound SA only the destination address of the outer IP header is used to identify the SA at the P-CSCF, together with the SPI. This address is the public address of the UE.

  - UE:
  For the outbound SA at the UE the source address shall be the local IP address of the UE, the destination address shall be the address of the P-CSCF as in the destination address of the IP header of the initial unprotected REGISTER message. For the inbound SA only the destination address of the outer IP header is used to identify the SA at the UE. This address shall be the IP address of the P-CSCF.

  Other data of the outer IP header (apart from IP addresses) shall be constructed as specified in [14].

- Ports used in the encapsulating UDP header:

  In case UDP encapsulated tunnel mode is selected, an encapsulating UDP header is inserted after the outer IP header. With respect to the ports used in the UDP header, the following rules shall be applied in accordance with standard [26] and [27]:

  - UE:
  Each protected and UDP encapsulated packet shall use port 4500 as source and destination port in the encapsulating UDP header.

  - P-CSCF:
  When the UE sends an UDP encapsulated packet towards the P-CSCF with the ports as described in the previous paragraph, the NAT will change the source port to a port different from 4500. This port is called port_Uenc. When the P-CSCF receives the first protected and UDP encapsulated message from the UE it shall store port_Uenc (cf. Section 7.2). From then on, all protected UDP encapsulated messages from the P-CSCF to the UE shall use port 4500 as source port and port_Uenc as destination port in the encapsulating UDP header.

The following rules apply:

1. For each unidirectional SA which has been established and has not expired, the SIP application at the P-CSCF stores at least the following data: (UE_IP_address, UE_protected_port, P-CSCF_protected_port, SPI, IMPI, IMPU1, ... , IMPUn, lifetime) in an "SA_table". The pair (UE_protected_port, P-CSCF_protected_port) equals either (*port_uc*, *port_ps*) or (*port_us*, *port_pc*).

NOTE: The SPI is only required when initiating and deleting SAs in the P-CSCF. The SPI is not exchanged between IPsec and the SIP layer for incoming or outgoing SIP messages.

2. The SIP application at the P-CSCF shall check upon receipt of a protected REGISTER message that the source IP address in the packet headers coincide with the UE's IP address inserted in the Via header of the protected REGISTER message. If the Via header does not explicitly contain the UE's IP address, but rather a symbolic name then the P-CSCF shall first resolve the symbolic name by suitable means to obtain an IP address.

3. The SIP application at the P-CSCF shall check upon receipt of an initial REGISTER message that the pair (UE_IP_address, UE_protected_client_port), where the UE_IP_address is the source IP address in the packet header and the protected client port is sent as part of the security mode set-up procedure (cf. clause 7.2), has not

yet been associated with entries in the "SA_table". Furthermore, the P-CSCF shall check that, for any one IMPI, no more than six SAs per direction are stored at any one time. If these checks are unsuccessful the registration is aborted and a suitable error message is sent to the UE.

NOTE: According to clause A.7.4 on SA handling, at most six SAs per direction may exist at a P-CSCF for one user at any one time.

In addition, if the P-CSCF detects that the UE is located behind a NAT (cf. Section 7.2), the P-CSCF shall check upon receipt of an initial REGISTER message that the triplet (UE_IP_address, UE_protected_client_port, UE_protected_server_port), where the UE_IP_address is the source IP address in the packet header and the protected client and server ports are sent as part of the security mode set-up procedure (cf. clause 7.2), has not yet been associated with entries in the "SA_table". If this check is unsuccessful the registration is aborted and a suitable error message is sent to the UE.

NOTE: The P-CSCF shall not accept registration attempts from UEs with the same address and protected server port in order to avoid ambiguities. Such situations may occur in case of multiple UEs behind the same NAT, which are assigned the same public IP address by the NAT.

4. For each incoming protected message the SIP application at the P-CSCF shall verify that the correct inbound SA according to clause A.7.4 on SA handling has been used. The SA is identified by the triple (UE_IP_address, UE_protected_port, P-CSCF_protected_port) in the "SA_table". The SIP application at the P-CSCF shall further check that the IMPU associated with the SA in the "SA_table" and the IMPU in the received SIP message coincide. If this is not the case the message shall be discarded.

5. For each unidirectional SA which has been established and has not expired, the SIP application at the UE stores at least the following data: (UE_protected_port, P-CSCF_protected_port, SPI, lifetime) in an "SA_table". The pair (UE_protected_port, P-CSCF_protected_port) equals either (*port_uc, port_ps*) or (*port_us, port_pc*).

NOTE: The SPI is only required to initiate and delete SAs in the UE. The SPI is not exchanged between IPsec and the SIP layer for incoming or outgoing SIP messages.

6. When establishing a new pair of SAs (cf. clause 6.3) the SIP application at the UE shall ensure that the selected numbers for the protected ports do not correspond to an entry in the "SA_table".

NOTE: Regarding the selection of the number of the protected port at the UE it is generally recommended that the UE randomly selects the number of the protected port from a sufficiently large set of numbers not yet allocated at the UE. This is to thwart a limited form of a Denial of Service attack. UMTS PS access link security also helps to thwart this attack.

7. For each incoming protected message the SIP application at the UE shall verify that the correct inbound SA according to clause A.7.4 on SA handling has been used. The SA is identified by the pair (UE_protected_port, P-CSCF_protected_port) in the "SA table".

NOTE: If the integrity check of a received packet fails then IPsec will automatically discard the packet.

A.7.2 Set-up of security associations (successful case)

The set-up of security associations is based on RFC 3329 [21]. Annex H of this specification shows how to use RFC 3329 [21] for the set-up of security associations.

In this clause the normal case is specified i.e. when no failures occurs. Note that for simplicity some of the nodes and messages have been omitted. Hence there are gaps in the numbering of messages, as the I-CSCF is omitted.

For the purpose of the description of the message processing in case UDP encapsulated tunnel mode is used, a conceptual functional element called "UDP encapsulation application" is used. The UDP encapsulation application handles all tasks relevant to the UDP encapsulation processing, i.e. the addition and removal of UDP headers to packets. In that sense it does not perform any IPSec processing as such. From an implementation point of view, it is immaterial whether the UDP encapsulation application and the IPSec processing are combined or kept separate. The UDP encapsulation application may reside on the P-CSCF or in a separate device.

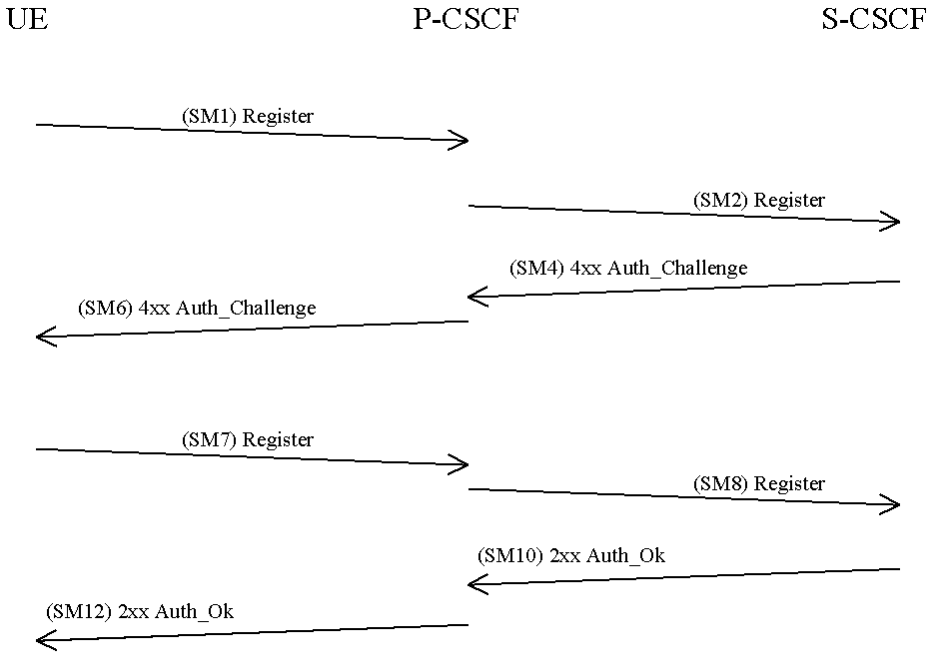


Figure 8

The UE sends a Register message towards the S-CSCF to register the location of the UE and to set-up the security mode, cf. clause A.6.1. In order to start the security mode set-up procedure, the UE shall include a *Security-setup*-line in this message.

The *Security-setup-line* in SM1 contains the Security Parameter Index values ~~and~~, the protected ports selected by the UE~~. It also contains~~ and a list of identifiers for the integrity and encryption algorithms, which the UE supports. It also contains the list of IPSec modes (i.e. transport or UDP encapsulated tunnel mode) supported by the UE.

SM1:
REGISTER(Security-setup = *SPI_U, Port_U, UE integrity and encryption algorithms list, IPSec mode list*)

SPI_U is the symbolic name of a pair of SPI values (cf. clause 7.1) (spi_uc, spi_us) that the UE selects. *spi_uc* is the SPI of the inbound SA at UE's the protected client port, and *spi_us* is the SPI of the inbound SA at the UE's protected server port. The syntax of spi_uc and spi_us are defined in Annex H.

Port_U is the symbolic name of a pair of port numbers (*port_uc, port_us*) as defined in clause 7.1. The syntax of port_uc and port_us is defined in Annex H.

~~Upon receipt of SM1, the P-CSCF temporarily stores the parameters received in the *Security-setup-line* together with the UE's IP address from the source IP address of the IP packet header, the IMPI and IMPU. Upon receipt of SM4, the P-CSCF adds the keys IK_{IM} and CK_{IM} received from the S-CSCF to the temporarily stored parameters.~~

~~A Release 6 P-CSCF shall propose SA alternatives for Release 5 and Release 6 UE's since the UE may or may not support confidentiality protection. The P-CSCF selects the SPI for the inbound SA. The P-CSCF then selects the SPIs for the inbound SAs. The same SPI number shall be used for Release 5 and Release 6 options. The P-CSCF shall define the SPIs such that they are unique and different from any SPIs as received in the *Security-setup-line* from the UE.~~

Editor's note: in this version of the document the NAT traversal of the unprotected messages is not described. As mechanism to allow for NAT traversal of unprotected messages, it is assumed that a SIP ALG is used on the P-CSCF. However, the functionality of such a SIP ALG has not been specified yet. It is expected that it will be specified in another 3G specification, which will also apply to the NAT traversal of the unprotected messages. It should be noted that it is assumed that the SIP ALG does not interfere with the SIP header fields with respect to the protected SIP messages.

Upon receipt of SM1, the P-CSCF temporarily stores the parameters received in the *Security-setup-line* together with the UE's IP address from the source IP address of the IP packet header, the IMPI and IMPU.

If the source IP address of the IP packet header is different from the address contained in the Via header, the P-CSCF adds a "received" parameter with the source IP address to the Via header following the rules of SIP message processing according to [6]. In this case the P-CSCF concludes that the UE is located behind a NAT device. If the UE has not signalled support for UDP encapsulated tunnel mode in message SM1 the P-CSCF shall silently discard the message and stop performing any further steps.

Otherwise, if the source IP address of SM1 matches the UE address in the Via header, the P-CSCF concludes that the UE is not located behind a NAT. The P-CSCF then continues with the set-up of security associations as specified in section 7.2, otherwise it continues as specified in this annex.

Upon receipt of SM4, the P-CSCF adds the keys IK_{IM} and CK_{IM} received from the S-CSCF to the temporarily stored parameters.

The P-CSCF then selects the SPIs for the inbound SAs. The P-CSCF shall define the SPIs such that they are unique and different from any SPIs as received in the *Security-setup-line* from the UE.

NOTE: This rule is needed since the UE and the P-CSCF use the same key for inbound and outbound traffic.

In order to determine the integrity and encryption algorithm the P-CSCF proceeds as follows: the P-CSCF has a list of integrity and encryption algorithms it supports, ordered by priority, cf. Annex H. ~~Release 6 algorithms shall have higher priority than Release 5 algorithms.~~The P-CSCF selects the first algorithm combination on its own list which is also supported by the UE.

The P-CSCF then establishes two new pairs of SAs in the local security association database.

In case the P-CSCF has discovered before that the UE is located behind a NAT, it informs the UDP encapsulation application about the IPSec SA data relevant for the UDP encapsulation process. This data consists of the IP source and destination addresses of the outer IP headers and the SPIs used in all four SAs established. At this point in time the UDP encapsulation application creates a table, the "UDP encapsulation table", with the following contents:

"UDP Encapsulation Table"				
	SA1	SA2	SA3	SA4
Src Addr	PCSCF	UE_pub	PCSCF	UE_pub
Dest Addr	UE_pub	PCSCF	UE_pub	PCSCF
Src Port	4500	***undef***	4500	***undef***
Dest Port	***undef***	4500	***undef***	4500
SPI	SPI_us	SPI_ps	SPI_uc	SPI_pc

The port 4500 shall be used as the source port for UDP encapsulated packets towards the UE and as the destination port for packets towards the P-CSCF. This is the IPSec standard port for UDP encpasulated IPSec packets (see [26],[27]). The source port for packets towards the P-CSCF and the destination port for packets towards the UE is not known yet and can only be learned in a later step (see below).

The *Security-setup*-line in SM6 contains the SPIs and the ports assigned by the P-CSCF. It also contains a list of identifiers for the integrity and encryption algorithms, which the P-CSCF supports. Furthermore, the P-CSCF indicates the IPSec mode of operation. In case the P-CSCF detected that the UE is behind a NAT, it indicates UDP encapsulated tunnel mode, otherwise transport mode is indicated.

NOTE: P-CSCF may be configured to trust on the encryption provided by the underlying access network. In this case, the P-CSCF acts according to Release 5 specificatons, and does not include encryption algorithms to the *Security-setup*-line in SM6.

SM6:
4xx Auth_Challenge(Security-setup = *SPI_P, Port_P, P-CSCF integrity and encryption algorithms list~~)~~, IPSec mode)*

SPI_P is the symbolic name of the pair of SPI values (cf. clause 7.1) (*spi_pc, spi_ps*) that the P-CSCF selects. *spi_pc* is the SPI of the inbound SA at the P-CSCF's protected client port, and *spi_ps* is the SPI of the inbound SA at the P-CSCF's protected server port. The syntax of spi_*pc* and *spi_ps* is defined in Annex H.

Port_P is the symbolic name of the port numbers (*port_pc, port_ps*) as defined in clause 7.1. The syntax of Port_P is defined in Annex H.

Upon receipt of SM6, the UE determines the integrity and encryption algorithms as follows: the UE selects the first integrity and encryption algorithm combination on the list received from the P-CSCF in SM 6 which is also supported by the UE.

NOTE: Release 5 UE will not support any encryption algorithms, and will choose the first Release 5 integrity algorithm on the list received from the P-CSCF in SM6.

The UE then proceeds to establish two new pairs of SAs in the local SAD. According to the IPSec mode included in SM6, the UE will either configure UDP encapsulated tunnel mode or transport mode. If transport mode is used the UE continues with the set-up of security associations as specified in section 7.2, otherwise it continues as specified in this annex.

The UE shall integrity and confidentiality protect SM7 and all following SIP messages. In case UDP encapsulation is required, all packets are in addition UDP encapsulated according to [27]. Furthermore the integrity and encryption algorithms list, *SPI_P*, and *Port_P* received in SM6, and *SPI_U, Port_U sent in SM1* shall be included:

SM7:
REGISTER(Security-setup = *SPI_U, Port_U, SPI_P, Port_P, P-CSCF integrity and encryption algorithms list)*

~~After receiving SM7 from the UE,~~If UDP encapsulated tunnel mode is used, the UE shall use the following addresses and ports in the various headers of message SM7:

SIP header:
In the Via and Contact header the UE shall use its public IP address and protected server port. The UE learns its public IP address by inspecting the received parameter in the Via header included in message SM6, in case such a parameter is present.

Editor's Note: it is not recommended and not deemed useful in case of UDP encapsulated tunnel mode that the UE uses a fully qualified domain name (FQDN) in its Contact or Via header. If FQDNs shall still be allowed, their use is for further study.

IP and UDP/TCP headers are used as specified in A.7.1.

If UDP encapsulated tunnel mode is applied, the UE shall start sending keep alive messages according to [27]. This ensures that the NAT binding is kept alive for the duration of the registration.

When SM 7 arrives at the P-CSCF it is at first processed by the UDP encapsulation application. The UDP encapsulation application can now learn port_Uenc, which the NAT has chosen for the UDP encapsulated packet. The UDP encapsulation application inserts this port in the UDP encapsulation table, so that the table is complete.

"UDP Encapsulation Table"				
	SA1	SA2	SA3	SA4

Src Addr	PCSCF	UE_pub	PCSCF	UE_pub
Dest Addr	UE_pub	PCSCF	UE_pub	PCSCF
Src Port	4500	***Port_Uenc***	4500	***Port_Uenc***
Dest Port	***Port_Uenc***	4500	***Port_Uenc***	4500
SPI	SPI_us	SPI_ps	SPI_uc	SPI_pc

The UDP encapsulation application removes the UDP header from the IP packet and hands it over to the IPSec processing.

After successful IPSec processing the SIP application in the P-CSCF shall check whether the integrity algorithms list, *SPI_P* and *Port_P* received in SM7 is identical with ~~the~~the corresponding parameters sent in SM6. It further checks whether *SPI_U* and *Port_U* received in SM7 are identical with those received in SM1. If these checks are not successful the registration procedure is aborted. The P-CSCF shall include in SM8 information to the S-CSCF that the received message from the UE was integrity protected as indicated in clause 6.1.5. The P-CSCF shall add this information to all subsequent REGISTER messages received from the UE that have successfully passed the integrity and confidentiality check in the P-CSCF.

SM8:
REGISTER(Integrity-Protection = *Successful, Confidentiality-Protection = Seccessful,* IMPI*)*

The P-CSCF finally sends SM12 to the UE. SM12 does not contain information specific to security mode setup (i.e. a Security-setup line), but with sending SM12 not indicating an error the P-CSCF confirms that security mode setup has been successful.

After receiving SM12 not indicating an error, the UE can assume the successful completion of the security-mode setup.

An example of how to make use of two pairs of unidirectional SAs is illustrated in the figure below with a set of example message exchanges protected by the respective IPsec SAs where the INVITE and following messages are assumed to be carried over TCP.

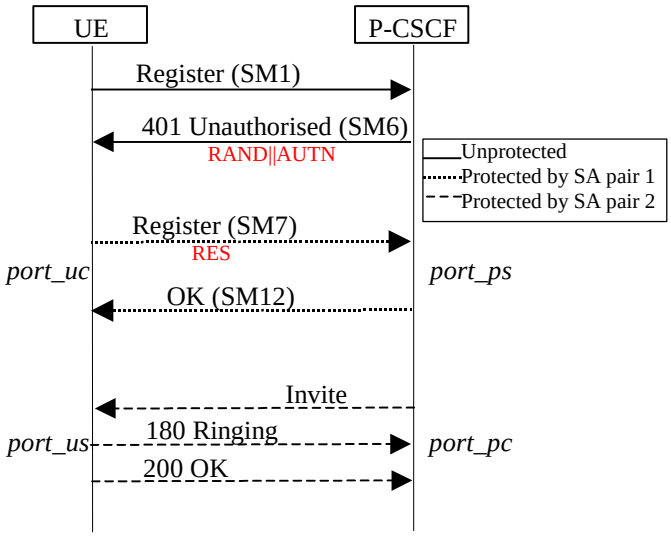


Figure 9

A.7.3 Error cases in the set-up of security associations

A.7.3.1 Error cases related to IMS AKA

~~Errors related to IMS AKA failures are specified in clause 6.1. However, this clause additionally describes how these shall be treated, related to security setup.~~

~~7.3.1.1 User authentication failure~~

~~In this case, SM7 fails integrity check by IPsec at the P-CSCF if the IK_{IM} derived from RAND at UE is wrong. The SIP application at the P-CSCF never receives SM7. It shall delete the temporarily stored SA parameters associated with this registration after a time-out.~~

~~In case IK_{IM} was derived correctly, but the response was wrong the authentication of the user fails at the S-CSCF due to an incorrect response. The S-CSCF shall send a 4xx Auth_Failure message to the UE, via the P-CSCF, which may pass through an already established SA. Afterwards, both, the UE and the P-CSCF shall delete the new SAs.~~

~~7.3.1.2 Network authentication failure~~

~~If the UE is not able to successfully authenticate the network, the UE shall send a REGISTER message which may pass through an already established SA, indicating a network authentication failure, to the P-CSCF. The P-CSCF deletes the new SAs after receiving this message.~~

~~7.3.1.3 Synchronisation failure~~

~~In this situation, the UE observes that the AUTN sent by the network in SM6 contains an out-of-range sequence number. The UE shall send a REGISTER message to the P-CSCF, which may pass through an already established SA, indicating the synchronization failure. The P-CSCF deletes the new SAs after receiving this message.~~

~~7.3.1.4 Incomplete authentication~~

~~If the UE responds to an authentication challenge from a S-CSCF, but does not receive a reply before the request times out, the UE shall start a registration procedure if it still requires any IM services. The first message in this registration should be protected with an SA created by a previous successful authentication if one exists.~~

~~When the P-CSCF receives a challenge from the S-CSCF and creates the corresponding SAs during a registration procedure, it shall delete any information relating to any previous registration procedure (including the SAs created during the previous registration procedure).~~

~~If the P-CSCF deletes a registration SA due to its lifetime being exceeded, the P-CSCF should delete any information relating to the registration procedure that created the SA.~~

The text in section 7.3.1 applies without changes.

A.7.3.2 Error cases related to the Security-Set-up

A.7.3.2.1 Proposal unacceptable to P-CSCF

In this case the P-CSCF cannot accept the proposal ~~set~~ sent by the UE in the Security-Set-up command of SM1. The P-CSCF shall respond to SM1 indicating a failure, by sending an error response to the UE.

A.7.3.2.2 Proposal unacceptable to UE

If the P-CSCF sends in the security-setup line of SM6 a proposal that is not acceptable for the UE, the UE shall abandon the registration procedure.

A.7.3.2.3 Failed consistency check of Security-Set-up lines at the P-CSCF

The P-CSCF shall check whether authentication and encryption algorithms list received in SM7 is identical with the authentication and encryption algorithms list sent in SM6. If this is not the case the registration procedure is aborted. (Cf. clause 7.2).

A.7.3.2.4 Missing NAT traversal capabilities in the presence of a NAT

In case the P-CSCF detects the presence of a NAT, but the UE or the P-CSCF do not support NAT traversal as specified in this annex, the P-CSCF shall abort the procedure.

A.7.4 Authenticated re-registration

Every registration that includes a user authentication attempt produces new security associations. If the authentication is successful, then these new security associations shall replace the previous ones. This clause describes how the UE and P-CSCF handle this replacement and which SAs to apply to which message.

When security associations are changed in an authenticated re-registration then the protected server ports at the UE (*port_us*) and the P-CSCF (*port_ps*) shall remain unchanged, while the protected client ports at the UE (*port_uc*) and the P-CSCF (*port_pc*) shall change. For the definition of these ports see clause 7.1.

If the UE has an already active pair of security associations, then it shall use this to protect the REGISTER message. If the S-CSCF is notified by the P-CSCF that the REGISTER message from the UE was integrity-protected it may decide not to authenticate the user by means of the AKA protocol. However, the UE may send unprotected REGISTER messages at any time. In this case, the S-CSCF shall authenticate the user by means of the AKA protocol. In particular, if the UE considers the SAs no longer active at the P-CSCF, e.g., after receiving no response to several protected messages, then the UE should send an unprotected REGISTER message.

Security associations may be unidirectional or bi-directional. This clause assumes that security associations are unidirectional, as this is the general case. For IP layer SAs, the lifetime mentioned in the following clauses is the lifetime held at the application layer. Furthermore deleting an SA means deleting the SA from both the application and IPsec layer. The message numbers, e.g. SM1, used in the following clauses relate to the message flow given in clause 6.1.1.

A.7.4.1 Void

A.7.4.1a Management of security associations in the UE

The UE shall be involved in only one registration procedure at a time, i.e. the UE shall remove any data relating to any previous incomplete registrations or authentications, including any SAs created by an incomplete authentication.

The UE may start a registration procedure with two existing pairs of SAs. These will be referred to as the old SAs. The authentication produces two pairs of new SAs. These new SAs shall not be used to protect non-authentication traffic until noted during the authentication flow. In the same way, certain messages in the authentication shall be protected with a particular SA. If the UE receives a message protected with the incorrect SA, it shall discard the message.

A successful authentication proceeds in the following steps:

- The UE sends the SM1 message to register with the IMS. If SM1 was protected, it shall be protected with the old outbound SA.
- The UE receives an authentication challenge in a message (SM6) from the P-CSCF. This message shall be protected with the old inbound SA if SM1 was protected and unprotected otherwise.
- If this message SM6 can be successfully processed by the UE, the UE creates the new SAs, which are derived according to clause 7.1. The lifetime of the new SAs shall be set to allow enough time to complete the registration procedure. If SM1 was protected and UDP encapsulated tunnel mode is used in the old SAs, the new SAs shall also be configured in with UDP encapsulated tunnel mode. The UE then sends its response (SM7) to the P-CSCF, which shall be protected with the new outbound SA. Meanwhile, if SM1 was protected, the UE shall use the old SAs for messages other than those in the authentication, until a successful message of new

authentication is received (SM12); if SM1 was unprotected, the UE is not allowed to use IMS service until it receives an authentication successful message (SM12).

- The UE receives an authentication successful message (SM12) from the P-CSCF. It shall be protected with the new inbound SA.

- After the successful processing of this message by the UE, the registration is complete. The UE sets the lifetime of the new SAs such that it either equals the latest lifetime of the old SAs or it will expire shortly after the registration timer in the message, depending which gives the SAs the longer life. For further SIP messages sent from UE, the new outbound SAs are used, with the following exception: when a SIP message is part of a pending SIP transaction it may still be sent over the old SA. A SIP transaction is called pending if it was started using an old SA. When a further SIP message protected with a new inbound SA is successfully received from the P-CSCF, then the old SAs shall be deleted as soon as either all pending SIP transactions have been completed, or have timed out. The old SAs shall be always deleted when the lifetime is expired. This completes the SA handling procedure for the UE.

A failure in the authentication can occur for several reasons. If the SM1 was not protected, then no protection shall be applied to the failure messages, except the user authentication failure message which shall be protected with the new SA. If SM1 was protected, the old SAs shall be used to protect the failure messages. In both cases, after processing the failure message, the UE shall delete the new SAs.

The UE shall monitor the expiry time of registrations without an authentication and if necessary increase the lifetime of the SAs created by the last successful authentication such that it will expire shortly after the registration timer in the message.

NOTE: In particular this means that the lifetime of a SA is never decreased.

The UE shall delete any SA whose lifetime is exceeded. The UE shall delete all SAs it holds once all the IMPUs are de-registered.

A.7.4.2 Void

A.7.4.2a Management of security associations in the P-CSCF

When the S-CSCF initiates an authentication by sending a challenge to the UE, the P-CSCF may already contain existing SAs from previously completed authentications. It may also contain two existing pairs of SAs from an incomplete authentication. These will be referred to as the old and registration SAs respectively. The authentication produces two pairs of new SAs. These new SAs shall not be used to protect non-authentication traffic until noted during the authentication flow. Similarly certain messages in the authentication shall be protected with a particular SA. If the P-CSCF receives a message protected with the incorrect SA, it shall discard the message.

The P-CSCF associates the IMPI given in the registration procedure and all the successfully registered IMPUs related to that IMPI to an SA.

A successful authentication proceeds in the following steps:

- The P-CSCF receives the SM1 message. If SM1 is protected, it shall be protected with the old inbound SA.

- The P-CSCF forwards the message containing the challenge (SM6) to the UE. This shall be protected with the old outbound SA, if SM1 was protected and unprotected otherwise.

- The P-CSCF then creates the new SAs, which are derived according to clause 7.1. The expiry time of the new SAs shall be set to allow enough time to complete the registration procedure. If SM1 was protected and UDP encapsulated tunnel mode was used in the old SAs, the new SAs shall also be configured with UDP encapsulated tunnel mode. The registration SAs shall be deleted if they exist.

- The P-CSCF receives the message carrying the response (SM7) from the UE. It shall be protected using the new inbound SA. If SM1 was protected, the old SAs are used to protect messages other than those in the authentication.

- The P-CSCF forwards the successful registration message (SM12) to the UE. It shall be protected using the new outbound SA. This completes the registration procedure for the P-CSCF. The P-CSCF sets the expiry time of the

new SAs such that they either equals the latest lifetime of the old SAs or it will expire shortly after the registration timer in the message, depending which gives the SAs the longer life.

- After SM12 is sent, the P-CSCF handles the UE related SAs according to following rules:
  - If there are old SAs, but SM1 belonging to the same registration procedure was received unprotected, the P-CSCF considers error cases happened, and assumes UE does not have those old SAs for use. In this case the P-CSCF shall remove the old SAs.
  - If SM1 belonging to the same registration procedure was protected with an old valid SA, the P-CSCF keeps this inbound SA and the corresponding three SAs created during the same registration with the UE active, and continues to use them. Any other old SAs are deleted. When the old SAs have only a short time left before expiring or a further SIP message protected with a new inbound SA is successfully received from the UE, the P-CSCF starts to use the new SAs for outbound messages with the following exception: when a SIP message is part of a pending SIP transaction it may still be sent over the old SA. A SIP transaction is called pending if it was started using an old SA. The old SAs are then deleted as soon as all pending SIP transactions have been completed, or have timed out. The old SAs are always deleted when the old SAs lifetime are expired. When the old SAs expire without a further SIP message protected by the new SAs, the new SAs are taken into use for outbound messages. This completes the SA handling procedure for the P-CSCF.

A failure in the authentication can occur for several reasons. If the SM1 was not protected, then no protection shall be applied to the failure messages, except the user authentication failure message which shall be protected with the new SAs. If SM1 was protected, the old SAs shall be used to protect the failure messages. In both cases, after processing the failure message, the P-CSCF shall delete the new SAs.

The P-CSCF shall monitor the expiry time of registrations without an authentication and if necessary increase the lifetime of SAs created by the last successful authentication such that it will expire shortly after the registration timer in the message.

The P-CSCF shall delete any SA whose lifetime is exceeded. The P-CSCF shall delete all SAs it holds that are associated with a particular IMPI once all the associated IMPUs are de-registered.

A.7.5 Rules for security association handling when the UE changes IP address

~~When a UE changes its IP address, e.g. by using the method described in RFC 3041 [18], then the UE shall delete the existing SA's and initiate an unprotected registration procedure using the new IP address as the source IP address in the packets carrying the REGISTER messages.~~

The text in section 7.5 applies without changes.

A.8 ISIM

The text in section 8 applies without changes.

***** END SET OF CHANGES *****

***** BEGIN SET OF CHANGES *****

Annex H (normative): The use of "Security Mechanism Agreement for SIP Sessions" [21] for security mode set-up

The BNF syntax of RFC 3329 [21] is defined for negotiating security associations for semi-manually keyed IPsec in the following way:

security-client	= "Security-Client" HCOLON sec-mechanism *(COMMA sec-mechanism)
security-server	= "Security-Server" HCOLON sec-mechanism *(COMMA sec-mechanism)
security-verify	= "Security-Verify" HCOLON sec-mechanism *(COMMA sec-mechanism)
sec-mechanism	= mechanism-name *(SEMI mech-parameters)
mechanism-name	= "ipsec-3gpp"
mech-parameters	= (preference / algorithm / protocol / mode / encrypt-algorithm / spi-c / spi-s / port-c / port-s)
preference	= "q" EQUAL qvalue
qvalue	= ("0" ["." 0*3DIGIT]) / ("1" ["." 0*3("0")])
algorithm	= "alg" EQUAL ("hmac-md5-96" / "hmac-sha-1-96")
protocol	= "prot" EQUAL ("ah" / "esp")
mode	= "mod" EQUAL ("trans" / "tun" / "UDP-enc-tun")
encrypt-algorithm	= "ealg" EQUAL ("des-ede3-cbc" /"aes-cbc" / "null")
spi-c	= "spi-c" EQUAL spivalue
spi-s	= "spi-s" EQUAL spivalue
spivalue	= 10DIGIT; 0 to 4294967295
port-c	= "port-c" EQUAL port
port-s	= "port-s" EQUAL port
port	= 1*DIGIT

The parameters described by the BNF above have the following semantics:

Mechanism-name: For manually keyed IPsec, this field includes the value "ipsec-3gpp". "ipsec-3gpp" mechanism extends the general negotiation procedure of RFC 3329 [21] in the following way:

1 The server shall store the Security-Client header received in the request before sending the response with the Security-Server header.

2 The client shall include the Security-Client header in the first protected request. In other words, the first protected request shall include both Security-Verify and Security-Client header fields.

3 The server shall check that the content of Security-Client headers received in previous steps (1 and 2) are the same.

Preference: As defined in RFC 3329 [21].

Algorithm: Defines the authentication algorithm. May have a value "hmac-md5-96" for algorithm defined in RFC 2403 [15], or "hmac-sha-1-96" for algorithm defined in RFC 2404 [16]. The algorithm parameter is mandatory.

Protocol: Defines the IPsec protocol. May have a value "ah" for RFC 2402 [19] and "esp" for RFC 2406 [13]. If no Protocol parameter is present, the value will be "esp".

NOTE: According to clause 6 only "esp" is allowed for use in IMS.

Mode: Defines the mode in which the IPsec protocol is used. May have a value "trans" for transport mode, and value "tun" for tunneling mode. If no Mode parameter is present, the value will be "trans".

NOTE: According to clause 6.3 ESP integrity shall be applied in transport mode i.e. only "trans" is allowed for use in IMS.

Encrypt-algorithm: If present, defines the encryption algorithm. May have a value "des-ede3-cbc" for algorithm defined in RFC 2451 [20] or "aes-cbc" for the algorithm defined in IETF RFC 3602 [22] or "null" if encryption is not used. If no Encrypt-algorithm parameter is present, the algorithm will be "null".

Spi-c: Defines the SPI number of the inbound SA at the protected client port.

Spi-s: Defines the SPI number of the inbound SA at the protected server port.

Port-c: Defines the protected client port.

Port-s: Defines the protected server port.

It is assumed that the underlying IPsec implementation supports selectors that allow all transport protocols supported by SIP to be protected with a single SA.

***** END SET OF CHANGES *****

Annex E: Improved IMS AKA for IPSec Traversal NAT

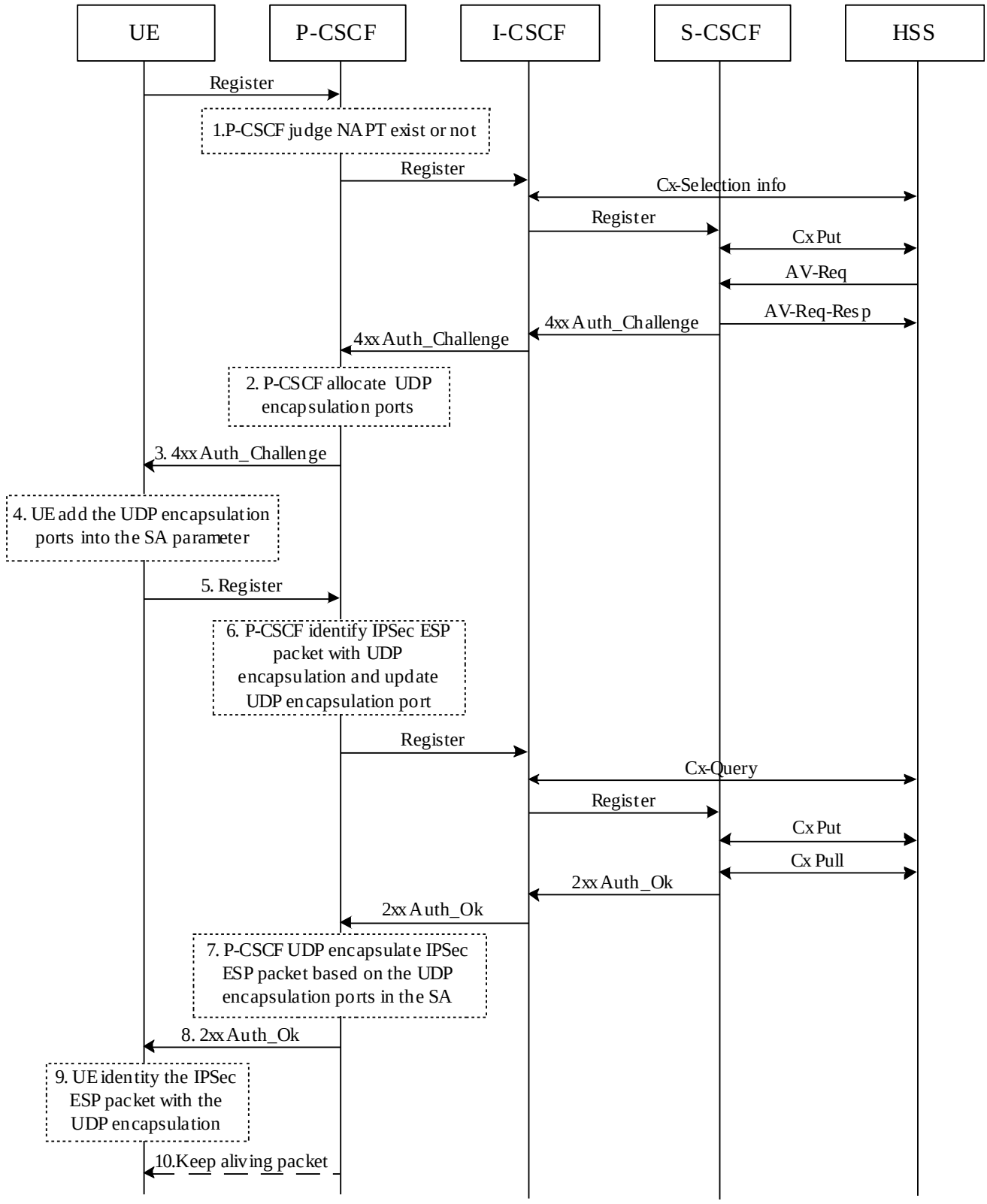
Editor's Note: Annex E is based on input document S3-050539 to SA3 #40.

E.1 Discussion

Because NAT is just friendly with TCP or UDP packet, IPSec ESP packet can only traverse NAT based on TCP or UDP encapsulation. And UDP encapsulation is preferable between UE and P-CSCF.

The document of S3-040720 has provided some information on how the SIP-AKA working with the NAT to implement the UDP encapsulation. It's a good idea for resolving the question about IPSec with NAT although there is some question on it. It's not reasonable that NAT mapping is created by the encapsulated ping packet, because SIP server A should record the source address and port after NAT, The Ping packet is not suitable.

Base on the idea of S3-040720, here introduced a improved IMS AKA to implement the UDP encapsulated IPSec ESP packet traversing NAT. The procedure is as following:



The difference between the above procedure with standard IMS AKA is a litter. just in UE and P-CSCF. Only the process in UE and P-CSCF has changed. The detail description is as following:

1. P-CSCF compare the source IP address of register packet with the UE IP address recorded in via header of register message. If it's same, there is no NAT between the UE and P-CSCF. Otherwise, NAT exists between UE and P-CSCF.

2. The P-CSCF extends two SA parameters to represent the UDP encapsulation source port and destination port. If there is no NAT, the UDP encapsulation ports will be zero or null. If there is a NAT, the P-CSCF will allocate two ports

as UDP encapsulation ports (Port_ues, Port_ued) and save them to SA. And the destination port can be same port number for all UE, such as 5061, the source port number can be different for each UE.

3. As the P-CSCF return the 4xx message to UE, the UDP encapsulation ports is also as SA parameter to send to UE. The message format is:

SM6:4xx Auth_Challenge(Security-setup = SPI_P, Port_P, Port_ues, Port_ued, P CSCF integrity and encryption algorithms list)

4. The UE also add the UDP encapsulation ports to SA. If the ports are zero or null, the UE will send IPSec ESP packet without UDP encapsulation to P-CSCF. Otherwise, the UE will send IPSec ESP packet with UDP encapsulation to P-CSCF in the following message.

5. If NAT exists, the IPSec ESP packet will be UDP encapsulated. And the UDP encapsulation ports will be also as message parameter and sent to P-CSCF, so the P-CSCF can check whether the ports are changed or not. The message format is:

SM7:REGISTER(Security-setup = SPI_U, Port_U, SPI_P, Port_P, Port_ues, Port_ued, P CSCF integrity and encryption algorithms list)

6. The P-CSCF identifies the IPSec ESP packet with UDP encapsulation by match the destination port of packet with the destination port (Port_ued) in the SA parameter. The matching will be simple if the Port_ued is same for all UE. If it's not matched, the UDP packet isn't IPSec ESP packet, otherwise it is, and maybe the source port (Port_ues) has been translated to Port_ues' by NAT, so the P-CSCF record the source port (Port_ues') of the packet and update source port in the SA parameter.

7. If the UDP encapsulation port in SA parameter isn't zero or null, the P-CSCF send the 2xx message to UE with UDP encapsulated IPSec ESP packet. And the UDP encapsulation ports should be reversed. The source and destination port in SA parameter should be the destination (Port_ues') and source (Port_ued) port of UDP packet.

8. If NAT exists, the IPSec ESP packet will be UDP encapsulated.

9. The UE identifies the IPSec ESP packet with UDP encapsulation by match the source (Port_ued) and destination (Port_ues) ports of packet with the destination (Port_ued) and source (Port_ues) ports in SA parameter. The compared ports are also reversed. If it's same, the UDP packet is IPSec ESP packet, otherwise, it's not.

10. After success registration and the UDP encapsulation ports in SA parameter is not zero or null, then the keep aliving packet will be sent periodic. It also can be sent from the UE.

The procedure extends two ports as UDP encapsulation ports in the SA parameter to resolve the IPSec traversing NAT question. The modification to IMS AKA is little. It's no impact with other function entity, including I/S-CSCF and HSS. And the implementation is simple.

Annex <X>:
Change history

Change history							
Date	TSG #	TSG Doc.	CR	Rev	Subject/Comment	Old	New
06-2005					*First version based on documents S3-050333, S3-050402, S3-050407, S3-040372 and S3-050427*		*0.0.1*
07-2005					*Changes based on drafting session at SA3 #39*	*0.0.1*	*0.0.2*
09-2005					*Changes based on agreements at SA3#40*	*0.0.2*	*0.1.0*
09-2005					*Changes based on comments to v0.1.0.*	*0.1.0*	*0.1.1*
11-2005					*New section 5.4 based on S3-050827 from SA3 #41*	*0.1.1*	*0.2.0*