3GPP TSG SA WG3 Security **S3-020092**

February 25 - February 28, 2002

Bristol, England

Agenda Item: 7.3

Source: Ericsson, Nortel Networks, Nokia

Title: Requirements and a proposed solution for SA_ID

Document for: Discussion and decision

1. Scope and objectives

This contribution address the issues related to the identifier of the Security Association established between the UE and the P-CSCF. It is currently not clarified in the TS 33.203 specification, how the SA_ID shall be used and what entity that shall handle the allocation of the SA_ID, in the case when SIP level protection is used between the UE and the P-CSCF.

A number of reasonable requirements on the SA_ID were listed in a contribution submitted to the ad-hoc meeting in Antwerp, however, there was no time to discuss the issue. This contribution updates the requirements and also proposes a solution.

It is proposed that SA3 adopts the requirements and the solution presented in this document as a working assumption. If such an agreement is made, the SIP layer integrity protection solution presented in 33.203 should be updated accordingly (see Appendix 1).

2 Requirements on SA_ID

In TS33.203 it is currently required that an SA_ID shall be used between the UE and the P-CSCF for identifying the algorithms and keys to be used. The Security Association between the UE and the P-CSCF is a fundamental element for SIP level integrity protection, as well as IPSec of course. It specifies what keys, algorithms etc that shall be used. According to Annex D the SPI shall be used for IPSec - one for each direction. Currently it has not been specified what requirements or how the SA_ID should look like for SIP-level protection.

2.1. Visibility

For IPSec the SA is one way or simplex between the nodes. Furthermore the SA in IPSec is security protocol specific and hence there will be a SA for each protocol (AH or ESP). In general, the SPI identifies the SA from other SAs to the same IP address and using the same security protocol and it is 32-bit long. The SPI is sent in every packet in clear between the nodes and the destination can use this value to fetch the correct SA. SPIs must be unique in the destination address, and consequently the receiver must issue its own SPIs. The SPI is re-used once the SA expires but it is guaranteed that the mapping <SPI, destination address, security protocol identifier> is one to one. In multihoming, also the source address is used to identify the SA.

Following the IPSec model it can be concluded that the SA_ID shall be sent in clear in each SIP message in both directions between the UE and the P-CSCF.

The most promising SIP layer solution for confidentiality protection is currently S/MIME with CMS. If SIP message is encrypted using S/MIME, the key identifier will be transferred in the CMS packet in clear. CMS does not set any special requirements for the identifiers of previously distributed symmetric keys. For example, the identifiers used in HTTP Digest could be used.

2.2. Uniqueness

It is clear that the P-CSCF is aware of the IP-address of the UE after the UE has registered. Assuming e.g. a terminating INVITE towards the UE which passes through the P-CSCF there is information at SIP level such that the P-CSCF is able to uniquely identify the UAS i.e. the UE. Hence this means that this piece of information shall have a one to one mapping to the SA_ID. Let us assume that this information is collected in the term Info. In theory Info could equal SA_ID but let us assume they are different but by knowing Info also the SA_ID is known. This is also valid for the IPSec alternative for IMS i.e. knowing Info at SIP level means that also the SPI is uniquely identified since no other information is available in the P-CSCF in this terminating scenario. Hence it follows that the SA_ID needs to be resolved at SIP level.

SIP layer protection will probably use HTTP Digest for integrity protection. The “integrity key” (i.e. the password) is identified by the client (i.e. the end-user) based on the information in the HTTP Digest challenge in *realm* parameter. The server or the proxy identifies the “integrity keys” based on the information in the HTTP Digest response in the *username* parameter. If HTTP Digest is used for integrity protection, the use of previous parameters for SA identification should be considered in order to guarantee the flexible development of the IMS security architecture. Furthermore, the solution for integrity protection should not prevent the use of HTTP Digest in the future; e.g. it must be possible to use HTTP Digest for authentication with application servers. In general, SA_ID needs to be unique but it is different for both directions in HTTP Digest.

2.3. Change of integrity keys

HTTP Digest does not identify different versions of integrity keys, and the concept of SA_ID is rather static. HTTP Digest is only using the keys that are valid “now”. This sets some additional requirements for key update. The update of integrity key must be secure, and there must be some mechanism to go back to use existing integrity keys if the key update is not successful.

3 Solution

This solution is based on HTTP Digest [rfc2617], the current SIP bis standard [sipbis7], and the integrity protection solution in 33.203 [33.203, S3z020042].

3.1 Parameters where SA-IDs are carried

SA-ID’s to identify the key to be used as HTTP Digest password are:

- UAC (both in UE and P-CSCF) must use the content of ‘realm’ parameter as a SA-ID.
- UAS (both in UE and P-CSCF) must use the content of ‘username’ parameter as a SA-ID. Optionally, UE may use the content of ‘responder’ parameter with extended HTTP Digest headers [digest-extensions], however, if both the ‘username’ and ‘responder’ parameters exist and they have different content, the ‘username’ should be used.

3.2 SA-IDs

SA-IDs for UE and P-CSCF are:

- UE in the role of UAC must put some user identifier, e.g. the IMPI, to the ‘username’ parameter.
- UE in the role of UAS must put the same use identifier as above, e.g. the IMPI, to the ‘realm’ parameter.
- P-CSCF in the role of UAC must put its globally unique realm to the ‘username’ parameter. Optionally, P-CSCF may put its globally unique realm to the ‘responder’ parameter with extended HTTP Digest headers [digest-extensions].
- P-CSCF in the role of UAS must put its globally unique realm to the ‘realm’ parameter.

In addition to the general rules above, the following rules must be fulfilled:

- The content of the realm parameter for P-CSCF must include some 3GPP specific key word, which is used by the UE to conclude that the IK will be valid for this realm. The use of additional semantics to the realm is explicitly allowed in [rfc2617]. For example, the realm for a P-CSCF in the Operator2 network could have a name "ik.p-cscf@operator2.com". The 3GPP specific key word for IK would be the "ik."-prefix in the URI. This kind of key word will guarantee flexible development of IMS architecture since the UE will use the key word in realm name for identifying the keys – and not trust that the one and only Proxy-Authenticate header comes from P-CSCF. There might be several such headers in this message in the future, and consequently it would be very difficult for UE to make the decision of the corresponding key.

- For terminating messages in which there are no 'realm' parameters, P-CSCF must use the Info parameter (discussed in chapter 2.2) to identify the 'realm' for terminating messages. (The final content of "Info" depends a lot on CN1, however, our current understanding is that "Info" will be the Request-URI modified by the S-CSCF to the terminating message.)

3.3 Change of integrity keys

The solution for key update is based on timers. When an authenticated re-registration takes place there will be two SAs available: the old SA (i.e. SA1) and the new SA (i.e. SA2). During the authenticated re-registration SM6 will reach the UE, which includes the challenge and the UE can check the authenticity of the message. The UE shall now in SM7 use the new SA i.e. SA2 and send the RES towards the S-CSCF. Upon receiving this message the P-CSCF expects that the UE shall use SA2 for protecting SIP messages to the P-CSCF and the UE expects to receive SM12 from the P-CSCF protected with the new SA2, if no time-out has been reached in the UE. In the successful case the old SA i.e. SA1 is deleted in the UE when the UE has received a SM12 from the P-CSCF protected with the new SA2. SA1 is deleted in the P-CSCF when the P-CSCF has received one additional message from the UE protected with the new SA2. If an expected SM12 message does not arrive before time-out in the UE or the P-CSCF does not receive an additional SIP message after SM7 from the UE, protected with the new SA2, the old SA i.e. SA1 is used in the UE and the P-CSCF. Hence the SA_ID can be re-used and the P-CSCF and the UE keeps track on the old and the new SA, which is a deterministic procedure.

Note that here might be a need for IMS specific rules on how the error situations are handled with HTTP Digest. HTTP Digest includes a mechanism for a server/proxy to communicate some information about the status of the username, password or nonce to the client. If a server/proxy adds a 'stale=true' parameter in an authentication challenge, the client will try using the same password (i.e. integrity key) with the delivered new nonce value. If the 'stale=false' or anything else, or if it is missing, the client must ask for a new password from the end-user. In IMS, stale values can be used to deal with different error situations related to the key update. For example, P-CSCF could ask the client to perform re-registration if it sent a "stale=false" parameter. The potential error situations are for further study.

4 Conclusions

- It is proposed that SA3 accept the following requirements applicable for the SA_ID, which is used in conjunction with HTTP Digest.

    1. The SA_ID shall be resolved at SIP level.

    2. The SA_ID shall be shared by the UE and the P-CSCF. It may be different or the same for both direction.

    3. The SA_ID value shall be unique in the UE and the P-CSCF in order to be able to distinguish between different UE's in the P-CSCF.

    4. The SA_ID shall be a static identifier associated with security data as security keys and algorithms, negotiated between the network and the UE. The UE and the P-CSCF shall keep track on the old and the new SA and apply them whenever no time-out is reached. If a time-out is reached the old SA shall be utilised.

    5. The SA_ID identifying the SA with the security keys, that is currently used to protect a SIP signalling message, shall be included as part of the SIP message. The SA_ID has to be sent in clear between the UE and the P-CSCF.

6. The SA_ID has to be included in all SIP signalling messages that are protected between the UE and the P-CSCF.

It is suggested that SA3 adopts the SA-ID solution presented in this document as a working assumption, and updates the SIP layer integrity protection solution in 33.203 accordingly (see Appendix 1).

References

[33.203] 3GPP "Access security for IP-based services", 3GPP TS 33.203, Release 5.

[digest-extensions] J. Undery et al. (2002) "SIP Digest Authentication: Extensions to HTTP Digest Authentication", IETF, Work in progress, January 2002.

[rfc2617] J. Franks et al. (1999) "HTTP Authentication: Basic and Digest Access Authentication", IETF, RFC 2617.

[S3z020042] Nortel Networks (2002) "Updates from IETF to SIP-Level Solution for IMS Integrity", 3GPP TSG SA WG3 Security, S3#21b, Tdoc S3z020042, 31st January - 1st February, 2002, Antwerp, Belgium.

[sipbis7] J. Rosenberg et al. (2002) "SIP: Session Initiation Protocol", IETF, Work in progress, February 2002, draft-ietf-sip-rfc2543bis-07.txt.

Appendix 1: Changes to the 33.203

C.2 6.3 Integrity mechanisms

[Editors note: There seems to be an unexpected shortcoming in the way SIP provides integrity protection on messages between UE and Proxies. In current SIP, HTTP Digest can be used to partially integrity protect the messages originated by an UE. However, SIP fails to provide integrity for Proxy to UE communication, i.e. for terminating INVITEs, for example. Proxies are not able to add Authorization headers on these messages, thus leaving the messages unprotected.

For the reason above, the headers and field names used in this section may not be final. However, the found inconsistency will probably make it easier for 3GPP to discuss about new SIP level integrity protection schemes with IETF.]

HTTP Digest shall provide integrity protection of SIP signalling between the UE and the P-CSCF, protecting all SIP signalling messages at the SIP level.

The SA that is required for Digest integrity protection shall use the 128-bit integrity key IK generated through IMS AKA, as specified in section 6.1. The integrity algorithm and key are identical for integrity protection applied to messages travelling in either direction. Negotiation of the integrity algorithm to use occurs in the following way: The UE communicates the set of integrity algorithms that it supports to the P-CSCF through the Security-setup header field of the REGISTER message, as described in section 7.2. The P-CSCF selects an algorithm to use from the set of algorithm capabilities common to both the UE and the P-CSCF. The P-CSCF indicates the algorithm to use in the "algorithm" directive of the Digest challenge that is subsequently issued to the UE.

Digest supports integrity protection of the SIP message body (not the headers) when the "qop-options" directive within the Digest challenge is set to the value "auth-int". Digest supports integrity protection of the SIP message body plus a named list of headers when the "qop-options" directive is set to the value "auth-hdr-int". Digest supports integrity protection of the entire SIP message when the "qop-options" directive within the Digest challenge is set to the value "auth-extd-int". (Use of either of these values of "qop-options" assumes that a context of client authentication has been previously established.) To provide for protection of the entire SIP message, the P-CSCF shall issue a Digest challenge to the UE specifying the value "auth-extd-int" for the "qop-options" directive.

The message 'digest', or message authentication code, is conveyed in the "response" directive of the Digest response. The rules for computing "response" are as described in [1] with the following consideration: if the UE receives a Digest challenge with the "~~qop-options~~realm" directive including a 3GPP specific key word (e.g. "ik.") ~~set to either "int" or "extended-intauth-extd-int", and the associated authentication challenge was an IMS AKA challenge~~, then the UE substitutes IK for the "password" component of A1 when computing "response=" in the Digest response. UE saves the content of the whole realm directive from the Proxy-Authentication header to be used as a key identifier for subsequent messages. At this stage UE can not be sure whether the proxy identified in the realm really knows the IK, however the Proxy-Authentication-Info header will be used for final verification. The UE sets the "username" component of A1 to some user identifier, e.g. the~~a~~ IMPI~~fixed value (e.g., "ims-user")~~. When sending messages to the UE that are to be integrity protected, the P-CSCF applies the same rules when computing "response". Within these terminating messages, the rules for the content of 'realm' and 'username' parameters are opposite than for originating messages: the "realm" directive will include the same user identifier as above, e.g. the IMPI, and the "username" the identifier of the P-CSCF (including the 3GPP specific key word, e.g. "ik."). In this manner, the whole SIP message is always protected.

Note that terminating messages arriving to the P-CSCF from the home network will probably not include IMPI. For these messages, P-CSCF must use some other identifier (e.g. Request-URI) to find the IMPI and the IK needed for the integrity protection.

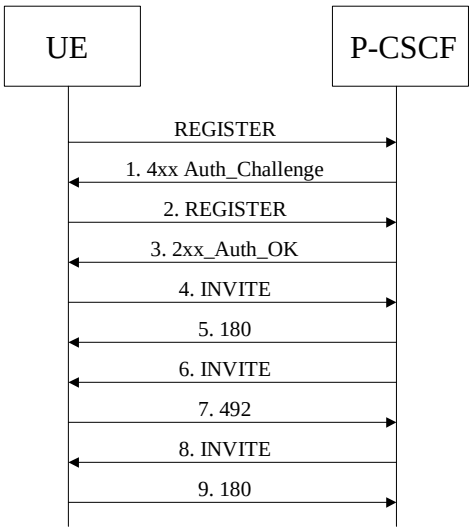
The Digest framework specifies that a server-initiated nonce is to be used by the client as a random number input to the production of the message digest. This nonce, along with a counter that is incremented by either endpoint when sending a message that is to be protected, facilitate anti-replay protection.

In the 3GPP IMS, then, normal operation of the Digest challenge-response mechanism for integrity protection is as follows:

Per RFC 2617, the Digest challenge-related directives are carried in either the WWW-Authenticate, or Proxy-Authenticate or UAS-Authenticate header fields. The P-CSCF adds a Proxy-Authenticate header field to the 4xx

Auth_Challenge that is sent by the S-CSCF (SIP registrar) toward the UE; the Proxy-Authenticate contains the Digest challenge that has been constructed by the P-CSCF.

Per RFC 2617, the Digest response-related directives are carried in either the Authorization, or Proxy-Authorization or UAS-Authorization header fields, depending upon which header field carried the corresponding Digest challenge. These directives contain the credentials for the message integrity check. In the IMS context, the UE responds to the initial Digest challenge by adding a Proxy-Authorization header field to the REGISTER toward the S-CSCF (registrar). The UE pre-emptively adds a Proxy-Authorization header field to all subsequent UE-initiated SIP requests. The UE and the P-CSCF adds the Proxy-Authentication-Info header to all SIP responses. **Finally, the P-CSCF adds an UAS-Authorization header field to all SIP requests sent toward the UE. Finally, the UE adds the UAS-Authentication-Info header to all SIP responses.** The simplified message flow shown below illustrates the relevant header fields and contents for the SIP-level integrity protection mechanism. Please note that the message flow contains three cases: a registration (1-3), and two SIP sessions: one UE initiated (4-5) and one UE terminated (6-9).



1. **4xx response – this carries both the IMS AKA challenge (from the registrar) and the Digest challenge for integrity protection (from the P-CSCF):**

   SIP/2.0 4xx Auth_Challenge

   WWW-Authenticate: EAP <RAND AUTN>

   Proxy-Authenticate: Digest realm=~~3GPP-IMS~~ik.p-cscf@operator2.com, nonce=<random-numberP-nonce1> algorithm=MD5 qop=extendedauth-extd-int

2. **Integrity protection is turned on with the next REGISTER – the integrity credentials are placed in the Digest response:**

   REGISTER sip: ... SIP/2.0

   Authorization: EAP <RES>

   Proxy-Authorization: Digest username=~~ims-user~~IMPI, realm=ik.p-cscf@operator2.com~~3GPP-IMS~~, nonce=<echo-random-numberP-nonce1>, uri=<SIP-URI>, response=<message-digest>, cnonce=<value>, nc=1, qop=extended-intauth-extd-int

3. **The 2xx response is also integrity protected – the P-CSCF adds the Proxy-Authentication-Info header to carry the message digest:**

   SIP/2.0 2xx Auth_Ok

   Proxy-Authentication-Info: Digest realm=ik.p-cscf@operator2.com, nextnonce=<P-nonce2>, qop=extended-intauth-extd-int, rspauth=<message-digest>, nc=21, cnonce=<value>

4. **A subsequent INVITE request must also be integrity protected – the UE pre-emptively adds the Proxy-Authorization header:**

INVITE sip: … SIP/2.0

Proxy-Authorization: Digest username=~~ims-user~~IMPI, realm=ik.p-cscf@operator2.com~~3GPP-IMS~~, nonce=<echo-random-numberP-nonce2>, uri=<SIP-URI>, response=<message-digest>, cnonce=<value>, nc=31, qop=extended-intauth-extd-int

Note: The client (UE) may re-use the previously issued nonce (i.e. set "nonce" to <P-nonce1> and "nc" to 1), but Digest recommends against this.

5. **The 180 is integrity protected in the same fashion was the 2xx response (message #3):**

SIP/2.0 180 Ringing

Proxy-Authentication-Info: Digest realm=ik.p-cscf@operator2.com, nextnonce=<P-nonce3>, qop=extended-intauth-extd-int, rspauth=<message-digest>, nc=41, cnonce=<value>

6. **An incoming INVITE must also be integrity protected – the first terminating SIP request, however, must be sent without the integrity credential (this permits the UE to issue a Digest challenge containing its own server-provided nonce).**

7. **The UE issues a 492 response containing a Digest challenge:**

SIP/2.0 492 Proxies Unauthorized

UAS-Authenticate: Digest realm=~~3GPP-IMS~~IMPI, nonce=<UE-nonce1>, algorithm=MD5, qop=auth-extd-int, target=ik.p-cscf@operator2.com~~<address>~~

8. **The P-CSCF adds the UAS-Authorization header, which has similar syntax to Proxy-Authorization:**

INVITE sip: … SIP/2.0

UAS-Authorization: Digest username=ik.p-cscf@operator2.com~~ims-user~~, realm=~~3GPP-IMS~~IMPI, nonce=<echo-random-numberUE-nonce1>, uri=<SIP-URI>, response=<message-digest>, cnonce=<value>, nc=51, qop=extended-intauth-extd-int, responder=ik.p-cscf@operator2.com~~<address>~~

9. **The UE protects the 180 response by adding UAS-Authentication-Info:**

SIP/2.0 180 Ringing

UAS-Authentication-Info: Digest realm=ik.p-cscf@operator2.com, nextnonce=<UE-nonce2>, qop=extended-intauth-extd-int, rspauth=<message-digest>, nc=61, cnonce=<value>

[Editors Note: Further details will be provided on how replay protection is accomplished. It has been identified that the scheme above needs to be enhanced since otherwise unnecessary loss of calls can occur. The reason for that is that both originating and terminating calls can occur and the counters in the P-CSCF and in the UE are not independent.]

[Editors Note: A description of the security mode setup headers shall be included in this Annex. Furthermore the message flows need to be enhanced.]

[Editors note: There might be a need for IMS specific rules on how the error situations are handled with HTTP Digest. HTTP Digest includes a mechanism for a server/proxy to communicate some information about the status of the username, password or nonce to the client. If a server/proxy adds a 'stale=true' parameter in an authentication challenge, the client will try using the same password (i.e. integrity key) with the delivered new nonce value. If the 'stale=false' or anything else, or if it is missing, the client must ask for a new password from the end-user. In IMS, stale values can be used to deal with different error situations related to the key update. For example, P-CSCF could ask the client to perform re-registration if it sent a "stale=false" parameter. The potential error situations are for further study.]