Secure Distribution of the Device Identity in Mobile Access Network

Konstantin Shemyak

Nokia Siemens Networks
konstantin.shemyak@nsn.com

Abstract. The paper presents an innovative way of providing cryptographic authentication credentials to mobile network elements. The proposed approach offers a practical solution to the problem of initial trust establishment between the newly installed hosts in the field and the existing network. It allows for true zero-touch secure start-up of the network elements.

Keywords: telecommunication network, network security, IP transport, device authentication, mobile IP backhaul.

1 Introduction

Telecommunication networks have traditionally used ATM lines for the data transport. Recently, mass availability and ease of the routing configuration of IP networks have made them an attractive option for telco operators.

ATM networks were originally deployed for commercial purposes and used private or semi-private fiber optic lines, and routers in controlled premises. Wiretapping on a fiber line is physically difficult, and if all switches are configured to prevent unauthorized management access, the network can be considered sufficiently secure (see [1], chapter 3, for elaboration on this topic).

In contrast, IP networks, although initially born as a military project, took off in the university world as free and open net. Such aspects as security, charging, or quality of service were not first goals of the initial design. They have been developed and standardized — in this order — generally later than the network itself. Standardization is still not always complete or perfect, as one can hear about, for example, "security gateway interoperability tests". The Internet is perceived by many users as something "insecure", requiring additional "protection", which is not available by default and for which many users are ready to pay.

With telecommunication networks, perception is different, as users expect, for example, that their phone calls can not be eavesdropped (at least easily), and do not generally expect an extra charge for something like "protection of my telephone line". Although this situation may change, it is a good idea for the telecom operators to take care that usage of the IP networks does not lower the overall security level.

A.U. Schmidt et al. (Eds.): MobiSec 2010, LNICST 47, pp. 117–126, 2010.

Technically, ways for reliable cryptographic protection of the IP traffic are known already for decades; reason for lower security lies probably in a fact that in many cases quick start and maximal flexibility are more important commercially than sufficient security level. Telecom operators must find the best balance between the two. This paper describes a solution, which simplifies the secure roll-out of the network, lowering cost of providing the secure IP network from the very beginning.

2 New Threats and Their Mitigation

Below are some of reasons why ATM lines are considered relatively safe when compared to IP lines:

- relatively high price of the equipment:
  Random example from year 2010: ATM router price is in range of 500-1000 euros, while Ethernet routers range under 100 euros.
- generally, a restrictive routing (approach "everything, which is not explicitly allowed, is denied" has been usually followed):
  In the IP world, notions like "hardening" or "firewall" appear sooner or later in almost any network. The out-of-box policy is not restrictive enough, so that some actions for security are needed **on top** of the main system. Exactly this allows for much faster network roll-outs, as there is no need to explicitly configure each route and rule for each packet: they are often "already there".
- less general popularity. This does not stop a deliberate paid intruder, but significantly lowers the entry barrier.

Main threats which the telco operator faces due to lower security of the transport network are eavesdropping, manipulation and impersonation of the traffic. In GSM and LTE networks, cryptographic protection of the radio interface terminates at the base station. Base stations are often located in a place without physical access control; especially smaller base stations are likely to be installed without any protective enclosures. Thus, unencrypted data becomes easily available on the wire. An attacker needs just minimal equipment (such as cheapest laptop) to get access to both user data and network control data. This can lead to disastrous consequences to the operator:

- Eavesdropping of user data. Intercepted conversation of any user can lead to legal penalties — or even more, if the user happens to be, for example, a high rank politician.
- Modifying or faking network management commands. This can have various impact depending on specific network; an easy-to-imagine attacker's action is to bring down part of the mobile network.

Besides the "main" threats listed above, there are others which are not affected by the network technology; to name a few: damaged or stolen equipment, modified software, and insider intrusions. We focus here only on the former ones, which become more probable with the shift to the IP mobile backhaul.

As noted in the introduction, technical means to protect the IP traffic are well known. Security protocols can be used at the network level (IPSec) or at the transport level (TLS). Both are standards both "de-jure" (standardized by IETF; see correspondingly [2] and [3]) and de-facto.

For any cryptographic protection, authentication of the remote peer is necessary, as otherwise it is trivial to organize a man-in-the-middle attack. Successful distribution of the authentication credentials completes the network security; options for authentication are considered in the next section.

3 Authentication: Standards and Open Points

With the endpoint authentication, both standards and real-world practices allow for some freedom. Common options for identification of the hosts in general IP networks are:

- pre-shared secrets
- RSA keys
- PGP keys
- X.509 certificates.

Pre-shared keys are not feasible for authentication in the network with nodes with potentially large numbers of peers. In WCDMA and LTE networks, the controller and the NodeB correspondingly can have tens or even hundreds of I_{ur} or $X1$ links for production traffic only. Nodes, carrying management traffic, can easily have thousands of peer network elements (think of a country-wide management center). As replacement of one host's key necessitates updates to all its peers peers, it is clear that management of solution, based on pre-shared keys, is not possible except if having same keys for all, or large groups of, hosts. The latter option is considered poor security, as a compromise of one key immediately leads to compromise of a whole network. Additionally, key distribution with the symmetrical credentials poses a problem by itself.

RSA keys, although being the asymmetric credentials and thus radically simplifying key distribution, do not allow for certifying one key via another, and thus pose the same scalability problem as pre-shared secrets.

PGP keys provide the most flexibility allowing arbitrary number of certifying keys for. But in practice they have not got wide acceptance for the purpose of *host* authentication (although working extremely well for *user* authentication in some scenarios).

X.509 certificates are by far the most common way of host authentication in the IP world. For example, virtually all secure sessions of web browsers are authenticated with X.509 certificates, issued by agencies which are trusted by the browser software.

In 2G and 3G networks, realizations of directing the data over IP links have not been covered by the standardization as something which belongs to an internal realization of a particular operator. In turn, for LTE networks, standards are present for protecting the backhaul link from the eNB to the core network

and for direct eNB-eNB communication. Usage of X.509 certificates for authentication of IPsec peers is specified in 3GPP TS 33.401 (see [4]) and detailed in TS 33.310 (see [5]). It is natural to expect that other network technologies (2G and 3G) will in practice follow the same line, because similar configuration at the nodes simplifies management of the security gateways serving multi-radio traffic.

The only remaining task in the authentication process is the distribution of the initial trust to the network elements in the field. A host, which is newly installed, or returned after repair, and the existing network must have knowledge about each other. Taking into account that the authentication is done with X.509 certificates, this means that each host needs to have trust in the certificate, presented by the peer. The next section describes possible approaches to this step.

4 Present and Discussed Approaches

Without any existing credentials, the only way for a reliable installation of the identity to a remote network peer is manual. An engineer shall copy the credentials — in practice, X.509 certificates — to the memory or a filesystem of the host. This has to be done only during the initial commissioning or after a repair. Normally, a technician is present at the site anyway for the mechanical installation; but this can be less skilled and less trusted person than the one managing the certificates. The latter is carrying sensitive information, which leak — for example, to a competitor or to a criminal — may have disastrous sequences for the operator.

Having reliable authentication credentials already out-of-the-box in the network elements would allow for cost savings and elimination of a potential weak link in the security chain. In scope of the mobile network, "reliable authentication credentials" means certificates, trusted by the network operator.

The problem which appears in this case is that the existing network, belonging to the operator, is usually not known to the equipment vendor at the manufacturing plant. Similarly, the new equipment piece is not known to the network peer of the new network element; locally, there is no trust to the certificate of the equipment vendor. Means to "familiarize" them with each other need to be devised. The following ways to do so can be listed:

1. Purchasing certificates from an existing external trust provider, which is trusted by both parties.
2. Obtaining certificates from some assigned common parent authority (this is different from the previous item in that the authority does not need to be "external"; it can be part of some organization common to the vendor, the operator, or both).
3. Cross-certification between the vendor and the operator.
4. Out-of-band delivery of the certificate information between the vendor and the operator.

We show our view on all of the listed approaches and give reasons for the selection of the last one.

4.1 Purchasing Certificates

Authentication with a certificate, purchased from an established "trust provider", is widely used on the Internet. This solution suits relatively well many Internet use cases when the secure identities of peers, such as a web service provider (for example a web shop) and a client, are not known in advance to each other. Such scheme, although, is not suitable for the case of network elements connecting to the operator's network, by the simple cost argument. Price for certificates from known providers is in range around 1000 euro (quote from VeriSign by year 2010, see [6]). This is not an acceptable price increase for a base station by orders of magnitude, even if volume discounts apply. Using a same certificate for all (or a large group of) network elements can not be considered secure enough by the same reason as using same pre-shared secret.

4.2 Using an Assigned Common Authority

Instead of a commercial trust provider, some organization with ties to both vendor and operator could provide mutually accepted trust anchors to both new network element and the operator's network. It can be, for example, some standardization body, such as 3GPP or GSMA.

Unfortunately, at the time of this writing (2010), such option can be considered only in future tense. Currently, no organization with sufficient trust, level of the technical preparation, and willingness to take such role, is visible on the horizon, while the secure identity provision is needed already now.

4.3 Cross-Certification

Vendor and operator can both issue a certificate to each other, and use them for signing certificates of end nodes. The end nodes process the certificate chains so that the trust chain from the peer certificate can be traced to a certificate, belonging to the own organization, and thus trusted. This can be a neat approach, but again it seems like belonging more to future than present. It would require certain level of experience with own PKI at the operator's end, and certain established relationships and processes between the vendor and the operator. Taking into account global nature of most vendors and operators, this seems to be a time-consuming task, requiring good level of synergy between the administrative and the technical bodies at both ends. Certainly being a viable candidate for future, this option was not considered possible for the networks which have to be deployed in year 2010.

4.4 Out-of-Band Delivery of the Identity

This solution works as pure "service" to the operator and does not require any PKI-related administrative activity between the vendor and the operator. It is based on providing to the operator a server, which receives information about vendor's identities. This server will authenticate the network element and request

the operator's identity to be installed, thus making the bind from the vendor's PKI to the operator's PKI.

In order to perform the latter task, this server needs some identification of the network element, which would be used at the operator's side (in practice, a convenient authentication is the hardware serial number). Such server received name "IDentity Mangement Server".

5 The Proposed Solution

5.1 Overview

Proposed system of the secure identity delivery is based on three steps:

1. Installation of the vendor-specific certificates at the factory
2. Delivery of the equipment information (in practice, it normally means serial numbers in this context) to the operator
3. Authenticating the base station at the operator's network with the vendor's specific Identity Management Server (IDM). This is a host owned by the network operator, but not necessarily logically belonging to the operator's network - it may reside in own demilitarized zone. This server is provided by the vendor. It connects to the operator's registration authority and retrieves a vendor certificate for the network element.

Below, each of these steps is described.

5.2 Installation of the Vendor's Certificates

A Registration Authority (RA) is established at each of the vendor's factories. When a network element (NE) is otherwise ready for shipment, it runs a collection of self-tests. A new step, which is not a test by itself but rather an extra production step, is added to this collection. During this step, the network element generates a certificate signing request (CSR) internally and contacts the RA for the signature. The link between the network element and the RA is in physically controlled premises and belongs only to the internal network (this is why the RA must be factory-specific). The private key of the certificate never leaves the internal storage of the network element. Two latter facts allow for the trust between RA and NE.

Factory-specific Registration Authority forwards the CSR to the vendor-specific Certificate Authority (CA). These two are connected by previously established secure link (VPN or SSL tunnel, maintained by the vendor); such links are normally present between any sites of any global manufacturer. Thus such CA does not need to reside in the same factory; in fact most natural choice for the vendor is to have just one CA for this purpose. The CA signs the request and returns it to the RA; then the RA installs it to the network element so that the latter has now the operator's identity.

At the same time, the CA records the fact that a particular network element has received a certificate. This record is stored in the equipment database together with the hardware information, such as the serial number.

5.3 Delivery of the Equipment Information

When the shipment of the network equipment is being performed, at the same time the database record indicating that a particular network element received a vendor's certificate, is sent to the operator. It can be done in a same data bundle as other hardware or licensing information, delivered to the customer. On the operator's end, the database record is received by the Identity Management Server, running the special vendor's software to receive such records.

5.4 Authenticating at the Operator's End

On the startup, IP hosts normally acquire own IP address via DHCP. In the operator's network, the DHCP server providing addresses to the new network elements delivers an additional parameter via the "vendor specific" field of the protocol, defined in RFC2132 ([7]). This parameter contains the IP address of the IDM.

After receiving and configuring these IP addresses (among with others, such as for example DNS server address), the base station establishes connection to the IDM server. The latter is the only entity in the operator's network, which knows about vendor's certificates and has trust in them. The server also has read access to the operator's database of the hardware information. Figure 1 shows overall architecture of the server.

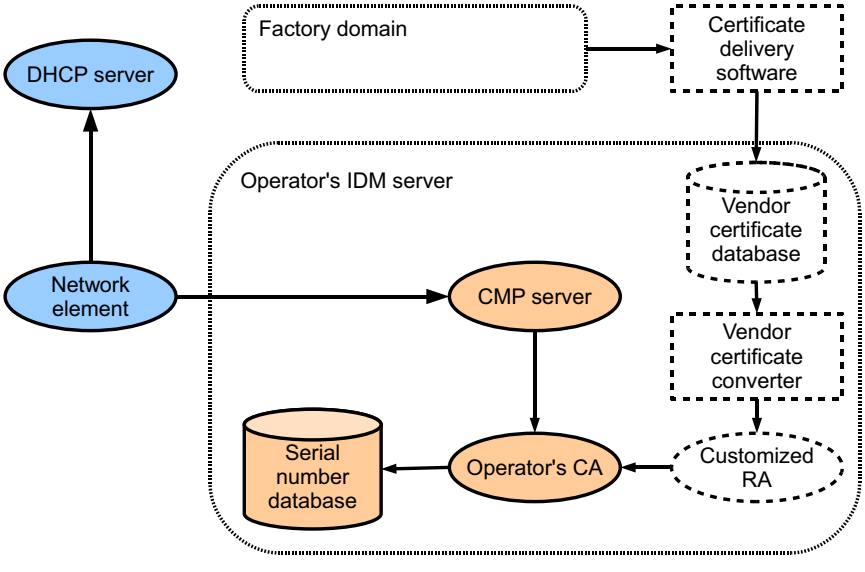

Fig. 1. Architecture of the operator's IDM server

At the IDM, the network element issues the key initialization request to the CMP server (which is the actual entity listening at the IP address, delivered via the DHCP message). NE authenticates with the vendor certificate. IDM server first verifies the certificate validity and the signature. Next, it checks in the operator's database that this network element is legitimate according to the identity recorded in the subject. The checks to be performed are up to the operator, but they may naturally include two points:

- the NE was actually sold to this particular operator, i.e. its serial number is present in the hardware database;
- the NE is installed to the correct area.

After checking the above, IDM server issues the certificate signing request to the operator's CA (or RA, depending on the operator's policy). The link between the IDM server and the CA/RA is secure, so the IDM server can authenticate towards the CA/RA. The latter, after issuing the operator's certificate, can do, at operator's discern, additional steps, such as for example bind the issued certificate with other properties of the network element stored at management databases.

The messaging flow of the network element enrollment is presented in Figure 2.

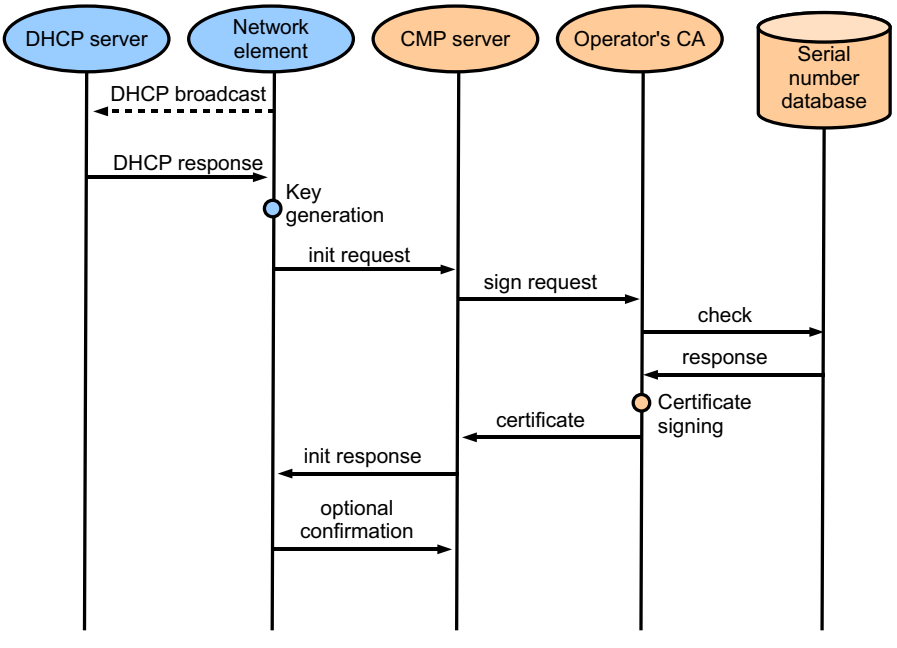

Fig. 2. Message flowchart of the network element enrollment

5.5 Decision on the Network Authentication

Note that at the first contact with the IDM, the network element does not authenticate the network. This can be seen as a limitation, but in fact it's a design feature. Authentication of the network eliminates only a denial of service-type threat: if a rogue network "catches" the network element, then the element connects to it and eventually the legitimate operator does not get this element to the own network. No other damage to the legitimate network can be done, as no un-authenticated hosts are accepted. In order to "catch" the network element, the rogue network must be on the same layer two link as the element itself in order to provide a rogue IP address, or to masquerade as the legitimate IP address. Similar effect can be achieved by, for example, flooding the access network, providing wrong routing information, faking DHCP response or just damaging the network element and/or its connections physically. These are denial-of-service types of attack, which are disturbing, but do not have as high impact as the active threats, and which can not be mitigated anyway without having full physical control over the access network. In case when such control is present, most of cryptographic protection of the traffic can be seen as redundant.

A base station, attached to the rogue network, can entice mobile devices to attach to it, but in 3G and 4G networks the user equipment will not proceed, as the radio network is authenticated by the handset. It is not in GSM networks, but there the problem of rogue base station has been known since the origination of the network and capturing of the existing legitimate base station does not allow for any advancement in this attack. For the operator, such attack will mean at most losing a newly installed base station until a maintenance visit.

From the cost viewpoint, authentication of the network will require knowledge of the future operator from the base station. Normally, operator is not known at the manufacturing phase; thus additional steps would be needed at the warehouse or at the customer service points. The latter increases costs of the manufacturing.

We conclude that the authentication of the network by the newly installed base station adds cost, but helps to mitigate only the threats which are present anyway and are not completely eliminated. This is the reason why the decision was taken to skip such authentication.

6 Conclusion

The paper gives an overview of methods for secure delivery of the identity for mobile network elements, which are also IP hosts. A practical and cost-effective solution is described. The approach is not limited by only case of mobile equipment vendor and mobile network operator, but is especially needed and most suitable namely in this case. It allows true zero-touch startup of the network element and is compatible with concepts of auto-connection. Need for administrative actions between the equipment vendor and the network operator is minimized.

The solution has been demonstrated at a Finnish data security conference "Tietoturvatapahtuma" ("Data Security Event") [8] in February 2010.

References

1. Tarman, T.D., Witzke, E.L.: Implementing security for ATM networks. Artech House, Inc., Boston (2002)
2. http://tools.ietf.org/html/rfc4301
3. http://tools.ietf.org/html/rfc5246
4. http://www.3gpp.org/ftp/Specs/html-info/33401.htm
5. http://www.3gpp.org/ftp/Specs/html-info/33310.htm
6. http://www.verisign.com/ssl/buy-ssl-certificates
7. http://www.ietf.org/rfc/rfc2132.txt
8. http://www.tietoturvatapahtuma.fi/default.htm