Authentication optimization for seamless handovers

Brahim Gaabab, David Binet, Jean-Marie Bonnin

Authentication Optimization for Seamless Handovers

Brahim Gaabab, *Graduate Student Member, IEEE*, David Binet, Jean-Marie Bonnin

***Abstract*— Handover is a key concept in wireless, cellular-based communications. It occurs when a mobile node changes its point of attachment to the network in order to meet an acceptable received signal strength level. Handover management refers to the procedures that need to be carried out upon handover, including discovery, configuration, authentication, and location update in order to recover network access. One important constraint to handover management is meeting the stringent real-time application requirements in terms of maximum allowed flow disruption. This paper addresses the authentication procedure, that is to say how the mobile node rapidly gets authenticated to restore the network access authorization? We focus on the authentication architecture being developed in the IETF including the recent Protocol for Carrying Network Access Authentication (PANA). We then analyze the different proposed optimizations and highlight their respective issues.**

***Index Terms*—Authentication, Handover Management, PANA.**

I. Introduction

Most of the current wireless communication technologies such as Universal Mobile Telecommunications System (UMTS), IEEE 802.11, and IEEE 802.16 are based on the cellular concept in order to provide an extensible service coverage. This means that several points of attachment (PoA) are installed to provide network access for nodes in their vicinity. Hence, while moving, a mobile node (MN) may need to change its PoA to the network in order to keep acceptable received signal strength. This event is called handover.

Wireless technologies provide their own mechanisms to deal with handover, in order for example to let the MN discover, select and reattach to PoAs. Nevertheless, a handover may have broader effects and span higher layers in the communication stack. A common example is *IP handover* [1], where the MN connects to a different IP subnet after handover and thus needs a new IP configuration, including an IP address and the default router address. In addition, some network-based states may require update such as routing and authorization data. On another hand, the handover management must cope with the stringent requirements imposed by real-time applications. The ITU-T recommends in [2] that handover latency, which is the interruption that occurs due the switching between PoAs, should not exceed 50 ms for such applications.

Authentication is a necessary procedure during a handover. It provides the network with the assurance about the mobile node identity to act accordingly. In this article, we focus on authentication for network access authorization. For this purpose we consider the Extensible Authentication Protocol (EAP) [3] as well as the Protocol for carrying Network Access Authentication (PANA) [4]. These protocols are designed in the Internet Engineering Task Force (IETF) and constitute a prospective architecture for network access authentication. Being part of the handover, the authentication process must not introduce significant delays in the handover latency and thus need to be optimized. The contribution of this paper is to analyze the effort done in this area in the IETF and highlights the issues that stems up in current solutions.

In the first section of the paper, we provide the reader with the necessary background including the EAP and PANA protocols. The second section introduces the two important optimizations in this field. These are analyzed in the third section and their issues are highlighted. The article ends by summarizing the challenges that still face standardization area.

II. Authentication and Authorization for Network Access

This section provides the required information about the authentication and authorization architecture for network access authorization that will serve for later analysis.

A. Extensible Authentication Protocol (EAP)

The Extensible Authentication Protocol (EAP) [3] was defined by the IETF to enable authentication between a client that requests network access and an authentication server. The latter is generally collocated with the operator's Authentication, Authorization and Accounting (AAA) server and thus we will use the terms interchangeably. The EAP architecture also introduces another entity called the *authenticator* in the signaling path between the authentication server and the client. The authenticator role is twofold; first it is the one that initiates EAP authentication process for the connecting clients. In addition, if ciphering keys -useful for encryption or integrity-protection of network access-are derived as result of a successful authentication, the authenticator ensures the distribution of those keys to the concerned entity.

EAP is actually a messaging protocol and does not

Brahim GAABAB is with France Telecom R&D, France, 42 rue de Coutures, BP 6243, 14066 Caen Cedex 4, and with Ecole Supérieure Nationale Supérieure de Bretagne, RSM Departement, France. (phone number: +33231759064, fax +33231735626; e-mail: brahim.gaabab@ orange-ft.com).

David BINET is with France Telecom R&D, France, 42 rue de Coutures, BP 6243, 14066 Caen Cedex 4. (e-mail: david.binet @ orange-ft.com).

Jean-Marie BONNIN is Ecole Nationale Supérieure de Bretagne, RSM Departement, France, (e-mail: jm.bonnin@enst-bretagne.fr).

implement authentication techniques. Instead, it is used to transport data units of *authentication methods* [3], such as the well known EAP-TLS [5], which supports client identification and credentials verification. In addition, in order to cope with the diversity of network access technologies and to ensure large deployment, EAP was designed impendent from the underlying environment; several EAP transport protocols have thus been proposed. Between the authentication server and the authenticator, which is in the operator's network, EAP is transported by RADIUS [6] or Diameter [7] protocols. On the access link (between the client and the authenticator), PPP and EAPoL (EAP over LAN) [8] were proposed for point to point and Ethernet network access technologies, respectively. In the next subsection, we describe a newly IETF designed protocol for transporting EAP over IP and which makes parts of the studied architecture.

B.Protocol for carrying Authentication for Network Access (PANA)

Key motivation for the design of the *Protocol for carrying Authentication for Network Access* (PANA) [4] was to support EAP authentication on top of IP access links. From this point of view, PANA provides an opportunity for operators to profit from IP abundance on access links and reduce management burden by reducing the number of used authentication transport protocols.

Besides encapsulation of EAP messages between the client and the authenticator (also called *PANA Authenticator*, PAA), PANA implements techniques to allow the PAA and the client discover each other before authentication starts. It also provides a scheme to derive session keys for securing network access.

The process of EAP/PANA authentication process is shown in Figure 1. After attaching to a network, the client first discovers the PAA, then asking for network access authorization. It also negotiates parameters that will be used for further key derivation. This is achieved in the handshake and discovery phase. Thereafter, it engages in an EAP authentication with the network and if it is successful, the AAA server determines authorization qualifications for the client and sends them to the authenticator using AAA protocol (cf. Figure 1, message 5). The authenticator also intervenes in authorization decision and can reject access request even if it is accepted by the AAA server due, for example, to a lack of bandwidth or whatever local resources. Finally, the authenticator notifies the client with the joint authorization decision (cf. Figure 1, message 6). It is just from that time on that the client can access the network.

In case authorization is granted, a state called *PANA session* is created at both client and the authenticator. Its main role is to store authorization parameters provided by the server, namely the authorization lifetime and the Master Session Key (MSK). Such key is used to bootstrap security at the access link (to apply for example a data encryption or integrity protection on link layer frames). The MSK is derived as a result of the EAP authentication at client and authentication server, and then transported to the authenticator within authentication success indication (cf. Figure 1, message 5).

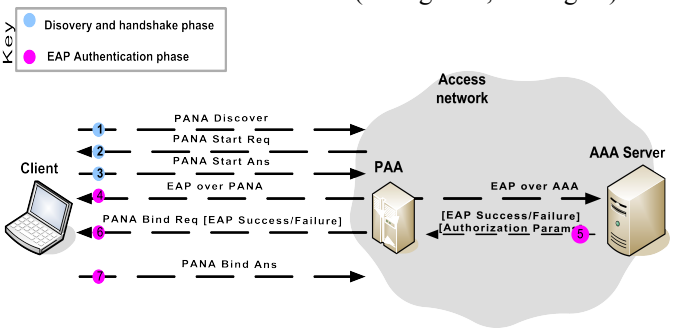

Figure 1. Authentication and Authorization Procedure

III.Authentication and Authorization Optimizations Review

There has been a great interest in optimizing the authentication and authorization procedure for the handover case. Upon handover, a mobile node may in fact find itself attached to a different PANA authenticator (PAA) than the one it has previously authenticated through. The newly discovered PAA has no state, that is a PANA session, to identify the mobile node and recognize its authorization rights, and therefore cannot authorize it to access the network. Actually, the mobile node has to carry out an EAP/PANA authentication through the new PAA to install a PANA session that stands for its authorization within it, and then recover its ongoing communication. However, such procedure (described in Figure 1) would last when using the EAP-TLS authentication method, no less than 6 round trips delay between the client and the authentication server. This largely over% the maximum allowed delay for real-time applications with for example a 100 ms round trip delay.

Hence one optimization problem we face in handover management field can be formulated as follows: how a mobile node obtains authorization and install a PANA session within a newly discovered PAA in a way that does not damage communications during a handover?

Two alternatives are proposed in the IETF PANA working group: the context transfer-based reauthorization and pre-authentication. In the following we briefly describe them and provide sufficient detail for their analysis and comparison in the next section.

A.Context Transfer Based Reauthorization

Contexts transfer [9]-[10] is a common paradigm in handover management research field. As its name suggests, it allows contexts re-location at the access network to fit mobile node movements without requiring long protocol operation for contexts installation. Context transfer can thus contribute to shorten handover latency [10].

This idea was jointly applied in [11]-[12] to shorten PANA session installation in a newly discovered authenticator by taking advantage from the existence of a previous one. The solution is shown in the Figure 2 and is described hereafter. We assume that the mobile node shares an unexpired PANA

session with previous PAA when it makes a handover. After discovering that it has switched to a new PAA, the mobile node transmits to it the old PANA session identifier as well as an authentication token (cf. Figure 2, messages 1 and 2), indicating that it wants a session transfer (cf. Figure 2, messages 3). The new PAA sends these data to the old one requesting PANA session transfer. The address of the previous PAA is determined from the session identifier which actually embeds authenticator's identity. The authentication token and the session identifier are used by the previous PAA to identify and authenticate the mobile node and then determine its authorization parameters. In case of successful authentication, the following parameters are sent to the new PAA: the remaining authorization lifetime and a key derived from the previous MSK, called *intermediate MSK*. Finally, the new PAA indicates successful authentication and authorization to the mobile node.

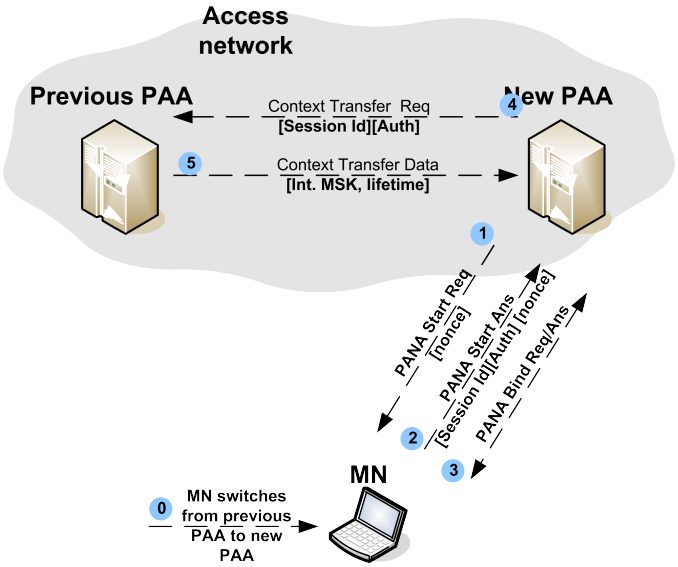

Figure 2. Context transfer based authorization upon handover

The new PANA session is constructed based on the two parameters received from the previous PAA. While the authorization lifetime is used as it is, the new MSK required for the new PANA session is derived from the security nonces (which are random numbers) exchanged during handshake phase between new PAA and the mobile node, as well as intermediate MSK.

B.Pre-authentication

Another paradigm in handover management research field is handover anticipation (also referred to as *make-before-break*). It consists in carrying out the procedures required for normal handover completion *before* it actually occurs; for instance a mobile node may like to obtain configuration for the target access router, install whatever context required to recover access, and even request for routing update to fit its movement.

Anticipation was proposed in [13] to allow a rapid PANA session installation at a target authenticator. Prior to pre-authentication, the mobile node have to discover the target authenticators it will switch to as result of a handover. Subsequently, it uses the original PANA exchange as described in Figure 1 through the current access network (yet with a slight modification to indicate that it is making pre-authentication). PANA messages are exchanged through the current point of attachment and are transparent to the local authenticator. If authorization is granted, a PANA session is derived at both mobile node and target PAA. It is kept inactive until the mobile node makes handover and notifies the PAA of its arrival through a request-answer exchange.

IV.Analysis of the Reauthorization Optimizations

Motivation for the design of the abovementioned solutions is to reduce the authentication and authorization contribution to the handover latency by rapidly installing the PANA session at the mobile node and the PAA. From this standpoint, pre-authentication clearly performs better than the context transfer based solution; after the handover, a mobile node would indeed just notify the PAA of its presence, whereas in the other solution, it takes no less than 4 round trips delay between the mobile node and the PAA with additional inter PAA communication.

Beyond the latency contribution, it is also important to determine at what cost optimizations are obtained. In the general case, in fact, optimizations may introduce additional assumptions that limit solution applicability or introduce new unwanted side effects. In the following paragraphs, we determine advantages and drawbacks of the presented solutions based on comparison to the original authentication and authorization architecture described in section II.

Concerning the context transfer based solution, two main differences are worth mentioned: the way authorization is decided on and the method for deriving the new PANA session parameters. In the discussed solution, a mobile node is authorized by solely ensuring that it was authorized within a previous authenticator, and that authorization lifetime has not expired. This violates the basic assumption that AAA server plays a central role in deciding authorization. For example, assume that an operator deploys both WiMax and WiFi access networks under separate authenticators. This operator offers WiFi-only access for nomadic access, and joint WiFi and WiMax access for seamless mobility. When a WiFi-only customer makes a handover to a WiMax access network, it should be prevented from access. Conversely, using context transfer would simply allow him access. Hence, to enable correct authorization decision, the AAA server which normally contains authorization qualifications of the customers must be involved.

Another difference to the original authentication and authorization procedure resides in the master session key (MSK) derivation. When making a context transfer, the MSK of the new authenticator is derived from the nonces exchanged during the handshake phase with the mobile node (cf. Figure 2, messages 1 and 2) as well as intermediate MSK transmitted by the old authenticator. An important security flaw resides in this

derivation scheme. Let's consider that an attacker has taken control over the old authenticator and discovered sufficient parameters to compute the intermediate MSK for a particular mobile node. If the attacker has in addition determined security nonces between the mobile node and the new authenticator, which are exchanged unencrypted, it can easily compute the new MSK. This violates one security requirement indicating that an MSK derived after client authentication through one authenticator must remain secret among both [14]. This would allow the attacker to determine the encryption and integrity protection keys and introduce further damages such as impersonation, sessions theft, etc.

Besides these differences, by introducing inter PAAs communication, the context transfer based solution requires that a mutual trust and an encrypted channel exist between PAAs. An immediate consequence is that the solution is not deployable for inter domain handovers as PAAs of different operators are not likely to have security relationship.

For pre-authentication, the main difference it presents compared to PANA resides in the use of anticipation. Recall that in pre-authentication, the mobile node has to discover the target authenticator before making a handover. However, as it cannot be sure of the actual handover target point of attachment it will switch to, it may end up with *several* authenticators as possible targets. This is due to the fact that generally the movement and the received radio signal are unpredictable. The mobile node conducts then a pre-authentication with each of them. Naturally, this inefficiently consumes network bandwidth, mobile node battery power, and storage capacity at those authenticators by introducing several inactive PANA sessions for just one mobile node. Anticipation also requires that mobile node and authenticator determine the optimal time to start pre-authentication. This impacts how much time the PANA session would be kept in the remote PAA in the inactive state. If pre-authentication is made long time before the handover takes place, PAA's capacity may be inefficiently limited and can prevent the addition of new PANA session of other mobile nodes. On the other hand, if it is made too late, the handover may occur abruptly leading to not installing the PANA session.

Another important issue concerns the authorization decision. Pre-authentication implicitly assumes that the authorization given in an anticipated manner would be the same as in normal authentication, i.e. PANA. This is not always true. In fact, the AAA server may base its authorization decision on the time at which it receives the authorization request as well as the dynamic state of its profile. As the exact handover time may not be determined nor controlled (it depends on the physical layer dynamics), the mobile node may become unauthorized just after handover due to a change for example in the profile.

However, an important feature of pre-authentication is that it does not introduce any assumptions on an existing relationship of both previous and new authenticators, as in context transfer solution.

V. Conclusions

This paper has discussed two IETF proposals to optimize the authentication and authorization procedure designed to cope with handover constraints. The analysis shows however that none of them meet the functional requirements related to security and authorization decision correctness. In addition, it shows that pre-authentication suffers additionally from the use of anticipation, which is not a well understood solution in the mobility area. From this standpoint, it seems that there is yet research to do in this area in order to provide new approaches for the authentication and authorization optimizations.

References

[1] D. Johnson, C. Perkins and J. Arkko, "Mobility Support in IPv6," RFC 3775, IETF Standards Track, June 2004.

[2] "Network performance objectives for IP-based services", ITU-T Recommendation Y.1541, May 2002.

[3] B. Aboba, L. Blunk, J. Vollbrecht, J. Carlson and H. Levkowetz, "Extensible Authentication Protocol (EAP)," RFC 3748, IETF Standards Track, June 2004.

[4] D. Forsberg, Y. Ohba, B. Patil, H. Tschofenig and A. Regin, "Protocol for Carrying Authentication for Network Access (PANA)," draft-ietf-pana-pana-12, August 2006, expires on February 23, 2007.

[5] B. Aboba and D. Simon, "PPP EAP-TLS Authentication Protocol," RFC 2716, IETF Experimental, October 1999.

[6] C. Rigney, S. Willens, A. Rubens and W. Simpson, "Remote Authentication Dial In User Service (RADIUS)," RFC 2865, IETF Draft Standard, June 2000.

[7] P. Calhoun, J. Loughney, E. Guttman, G. Zorn and J. Arkko, "Diameter Base Protocol," RFC 3588, IETF Standards Track, September 2003.

[8] "IEEE Standard for Local and metropolitan area networks – Port-Based Network Access Control," IEEE Std. 802.1X-2001, June 2001.

[9] J. Loughney, M. Nakhjiri, C. Perkins and R. Koodli, "Context Transfer Protocol (CxTP)," RFC 4067, IETF Experimental, July 2005.

[10] Koodli, R. & Perkins, C.E. "Fast Handovers and Context Transfers in Mobile Networks" *SIGCOMM Comput. Commun. Rev.*, ACM Press, 2001, vol. 31, pp. 37-47.

[11] D. Forsberg, Y. Ohba, B. Patil, H. Tschofenig and A. Yegin, "PANA Mobility Optimizations," draft-ietf-pana-mobopts-01, October 2005, expired on April 24, 2006.

[12] J. Bournelle, M. Laurent-Maknavicius, H. Tschofenig, Y. El Mghazli, G. Giaretta, R. Lopez and Y. Ohba, "Use of Context Transfer Protocol (CxTP) for PANA," draft-ietf-pana-cxtp-00, March 2006, expires on September 7, 2006.

[13] Y. Ohba, "Pre-authentication Support for PANA," draft-ohba-pana-preauth-01, March 2006, expires on September 4, 2006.

[14] R. Housley and Bernard Aboba, "Guidance to AAA Key Management," draft-housley-aaa-key-mgmt-03, July 2006, expires on February 2007.