Extensible Authentication Protocol (EAP)
Early Authentication Problem Statement

Abstract

Extensible Authentication Protocol (EAP) early authentication may be defined as the use of EAP by a mobile device to establish authenticated keying material on a target attachment point prior to its arrival. This document discusses the EAP early authentication problem in detail.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc5836.
1. Introduction

When a mobile device, during an active communication session, moves from one access network to another and changes its attachment point, the session may be subjected to disruption of service due to the delay associated with the handover operation. The performance requirements of a real-time application will vary based on the type of application and its characteristics such as delay and packet-loss tolerance. For Voice over IP applications, ITU-T G.114 [ITU] recommends a steady-state end-to-end delay of 150 ms as the upper limit and rates 400 ms as generally unacceptable delay. Similarly, a streaming application has tolerable packet-error rates ranging from 0.1 to 0.00001 with a transfer delay of less than 300 ms. Any help that an optimized handoff mechanism can provide toward meeting these objectives is useful. The ultimate objective is to achieve seamless handover with low latency, even when handover is between different link technologies or between different Authentication, Authorization, and Accounting (AAA) realms.
As a mobile device goes through a handover process, it is subjected to delay because of the rebinding of its association at or across several layers of the protocol stack and because of the additional round trips needed for a new EAP exchange. Delays incurred within each protocol layer affect the ongoing multimedia application and data traffic within the client [WCM].

The handover process often requires authentication and authorization for acquisition or modification of resources assigned to the mobile device. In most cases, these authentications and authorizations require interaction with a central authority in a realm. In some cases, the central authority may be distant from the mobile device. The delay introduced due to such an authentication and authorization procedure adds to the handover latency and consequently affects ongoing application sessions [MQ7]. The discussion in this document is focused on mitigating delay due to EAP authentication.

2. Terminology

AAA

Authentication, Authorization, and Accounting (see below). RADIUS [RFC2865] and Diameter [RFC3588] are examples of AAA protocols defined in the IETF.

AAA realm

The set of access networks within the scope of a specific AAA server. Thus, if a mobile device moves from one attachment point to another within the same AAA realm, it continues to be served by the same AAA server.

Accounting

The act of collecting information on resource usage for the purpose of trend analysis, auditing, billing, or cost allocation [RFC2989].

Attachment Point

A device, such as a wireless access point, that serves as a gateway between access clients and a network. In the context of this document, an attachment point must also support EAP authenticator functionality and may act as a AAA client.

Authentication

The act of verifying a claimed identity, in the form of a preexisting label from a mutually known name space, as the originator of a message (message authentication) or as the end-point of a channel (entity authentication) [RFC2989].
Authenticator
The end of the link initiating EAP authentication [RFC3748].

Authorization
The act of determining if a particular right, such as access to some resource, can be granted to the presenter of a particular credential [RFC2989].

Candidate Access Network
An access network that can potentially become the target access network for a mobile device. Multiple access networks may be candidates simultaneously.

Candidate Attachment Point (CAP)
An attachment point that can potentially become the target attachment point for a mobile device. Multiple attachment points may be candidates simultaneously.

Candidate Authenticator (CA)
The EAP authenticator on the CAP.

EAP Server
The entity that terminates the EAP authentication method with the peer [RFC3748]. EAP servers are often, but not necessarily, co-located with AAA servers, using a AAA protocol to communicate with remote pass-through authenticators.

Inter-AAA-realm Handover (Inter-realm Handover)
A handover across multiple AAA realms.

Inter-Technology Handover
A handover across different link-layer technologies.

Intra-AAA-realm Handover (Intra-realm Handover)
A handover within the same AAA realm. Intra-AAA-realm handover includes a handover across different authenticators within the same AAA realm.

Intra-Technology Handover
A handover within the same link-layer technology.

Master Session Key (MSK)
Keying material that is derived between the EAP peer and server and exported by the EAP method [RFC3748].

Peer
The entity that responds to the authenticator and requires authentication [RFC3748].
Serving Access Network
   An access network that is currently serving the mobile device.

Serving Attachment Point (SAP)
   An attachment point that is currently serving the mobile device.

Target Access Network
   An access network that has been selected to be the new serving access network for a mobile device.

Target Attachment Point (TAP)
   An attachment point that has been selected to be the new SAP for a mobile device.

3. Problem Statement

   The basic mechanism of handover is a two-step procedure involving
   o handover preparation and
   o handover execution

3.1. Handover Preparation

   Handover preparation includes the discovery of candidate attachment points and selection of an appropriate target attachment point from the candidate set. Handover preparation is outside the scope of this document.

3.2. Handover Execution

   Handover execution consists of setting up Layer 2 (L2) and Layer 3 (L3) connectivity with the TAP. Currently, handover execution includes network access authentication and authorization performed directly with the target network; this may include full EAP authentication in the absence of any particular optimization for handover key management. Following a successful EAP authentication, a secure association procedure is typically performed between the mobile device and the TAP to derive a new set of link-layer encryption keys from EAP keying material such as the MSK. The handover latency introduced by full EAP authentication has proven to be higher than that which is acceptable for real-time application scenarios [MQ7]; hence, reduction in handover latency due to EAP is a necessary objective for such scenarios.
3.2.1. Examples

3.2.1.1. IEEE 802.11

In IEEE 802.11 Wireless Local Area Networks (WLANs) [IEEE.802-11.2007] network access authentication and authorization involves performing a new IEEE 802.1X [IEEE.802-1X.2004] message exchange with the authenticator in the TAP to execute an EAP exchange with the authentication server [WPA]. There has been some optimization work undertaken by the IEEE, but these efforts have been scoped to IEEE link-layer technologies; for example, the work done in the IEEE 802.11f [IEEE.802-11F.2003] and 802.11r [IEEE.802-11R.2008] Task Groups applies only to intra-technology handovers.

3.2.1.2. 3GPP TS33.402

The Third Generation Partnership Project (3GPP) Technical Specification 33.402 [TS33.402] defines the authentication and key management procedures performed during interworking between non-3GPP access networks and the Evolved Packet System (EPS). Network access authentication and authorization happens after the L2 connection is established between the mobile device and a non-3GPP target access network, and involves an EAP exchange between the mobile device and the 3GPP AAA server via the non-3GPP target access network. These procedures are not really independent of link technology, since they assume either that the authenticator lies in the EPS network or that separate authentications are performed in the access network and then in the EPS network.

3.3. Solution Space

As the examples in the preceding sections illustrate, a solution is needed to enable EAP early authentication for inter-AAA-realm handovers and inter-technology handovers. A search for solutions at the IP level may offer the necessary technology independence.

Optimized solutions for secure inter-authenticator handovers can be seen either as security context transfer (e.g., using the EAP Extensions for EAP Re-authentication Protocol (ERP)) [RFC5296], or as EAP early authentication.

3.3.1. Context Transfer

Security context transfer involves transfer of reusable key context to the TAP and can take two forms: horizontal and vertical.
Horizontal security context transfer (e.g., from SAP to TAP) is not recommended because of the possibility that the compromise of one attachment point might lead to the compromise of another (the so-called domino effect, [RFC4962]). Vertical context transfer is similar to the initial establishment of keying material on an attachment point in that the keys are sent from a trusted server to the TAP as a direct result of a successful authentication. ERP specifies vertical context transfer using existing EAP keying material obtained from the home AAA server during the initial authentication. A cryptographically independent re-authentication key is derived and transmitted to the TAP as a result of successful ERP authentication. This reduces handover delay for intra-realm handovers by eliminating the need to run full EAP authentication with the home EAP server.

However, in the case of inter-realm handover, either ERP is not applicable or an additional optimization mechanism is needed to establish a key on the TAP.

3.3.2. Early Authentication

In EAP early authentication, AAA-based authentication and authorization for a CAP is performed while ongoing data communication is in progress via the serving access network, the goal being to complete AAA signaling for EAP before the mobile device moves. The applicability of EAP early authentication is limited to the scenarios where candidate authenticators can be discovered and an accurate prediction of movement can be easily made. In addition, the effectiveness of EAP early authentication may be less significant for particular inter-technology-handover scenarios where simultaneous use of multiple technologies is not a major concern.

There are also several AAA issues related to EAP early authentication, discussed in Section 8.

4. System Overview

Figure 1 shows the functional elements that are related to EAP early authentication. These functional elements include a mobile device, a SAP, a CAP, and one or more AAA and EAP servers; for the sake of convenience, the AAA and EAP servers are represented as being co-located. When the SAP and CAP belong to different AAA realms, the CAP may require a different set of user credentials than those used by the peer when authenticating to the SAP. Alternatively, the CAP and the SAP may rely on the same AAA server, located in the home realm of the mobile device (MD).
A mobile device is attached to the serving access network. Before the MD performs handover from the serving access network to a candidate access network, it performs EAP early authentication with a candidate authenticator via the serving access network. The peer may perform EAP early authentication with one or more candidate authenticators. It is assumed that each attachment point has an IP address. It is assumed that there is at least one CAP in each candidate access network. The serving and candidate access networks may use different link-layer technologies.

Each authenticator is either a standalone authenticator or a pass-through authenticator [RFC3748]. When an authenticator acts as a standalone authenticator, it also has the functionality of an EAP server. When an authenticator acts as a pass-through authenticator, it communicates with the EAP server, typically using a AAA transport protocol such as RADIUS [RFC2865] or Diameter [RFC3588].

If the CAP uses an MSK [RFC5247] for generating lower-layer ciphering keys, EAP early authentication is used to proactively generate an MSK for the CAP.

5. Topological Classification of Handover Scenarios

The complexity of the authentication and authorization part of handover depends on whether it involves a change in EAP server. Consider first the case where the authenticators operate in pass-through mode, so that the EAP server is co-located with a AAA server. Then, there is a strict hierarchy of complexity, as follows:

1. inter-attachment-point handover with common AAA server: the CAP and SAP are different entities, but the AAA server is the same. There are two sub-cases here:
   
   (a) the AAA server is common because both attachment points lie within the same network, or
(b) the AAA server is common because AAA entities in the serving
and candidate networks proxy to a AAA server in the home
realm.

2. inter-AAA-realm handover: the CAP and SAP are different entities,
and the respective AAA servers also differ. As a result,
authentication in the candidate network requires a second set of
user credentials.

A third case is where one or both authenticators are co-located with
an EAP server. This has some of the characteristics of an inter-AAA-
realm handover, but offers less flexibility for resolution of the
early authentication problem.

Orthogonally to this classification, one can distinguish intra-
technology handover from inter-technology handover thinking of the
link technologies involved. In the inter-technology case, it is
highly probable that the authenticators will differ. The most likely
cases are 1(b) or 2 in the above list.

6. Models of Early Authentication

As noted in Section 3, there are cases where early authentication is
applicable while ERP does not work. This section concentrates on
providing some models around which we can build our analysis of the
EAP early authentication problem. Different usage models can be
defined depending on whether

- the SAP is not involved in early authentication (direct pre-
  authentication usage model),
- the SAP interacts only with the CAP (indirect pre-authentication
  usage model), or
- the SAP interacts with the AAA server (the authenticated
  anticipatory keying usage model).

It is assumed that the CAP and SAP are different entities. It is
further assumed in describing these models that there is no direct L2
connectivity between the peer and the candidate attachment point.

6.1. EAP Pre-Authentication Usage Models

In the EAP pre-authentication model, the SAP does not interact with
the AAA server directly. Depending on how the SAP is involved in the
pre-authentication signaling, the EAP pre-authentication usage model
can be further categorized into the following two sub-models, direct
and indirect.
6.1.1. The Direct Pre-Authentication Model

In this model, the SAP is not involved in the EAP exchange and only forwards the EAP pre-authentication traffic as it would any other data traffic. The direct pre-authentication model is based on the assumption that the MD can discover candidate authenticators and establish direct IP communication with them. It is applicable to any of the cases described in Section 5.

![Figure 2: Direct Pre-Authentication Usage Model](image)

The direct pre-authentication signaling for the usage model is shown in Figure 3.

![Figure 3: Direct Pre-Authentication Signaling for the Usage Model](image)

6.1.2. The Indirect Pre-Authentication Usage Model

The indirect pre-authentication usage model is illustrated in Figure 4.
In the indirect pre-authentication model, it is assumed that a trust relationship exists between the serving network (or serving AAA realm) and candidate network (or candidate AAA realm). The SAP is involved in EAP pre-authentication signaling. This pre-authentication model is needed if the peer cannot discover the candidate authenticators identity or if direct IP communication between the MD and CAP is not possible due to security or network topology issues.

The role of the SAP in this pre-authentication model is to forward EAP pre-authentication signaling between the mobile device and CAP; the role of the CAP is to forward EAP pre-authentication signaling between the peer (via the SAP) and EAP server and receive the transported keying material.

The pre-authentication signaling for this model is shown in Figure 5.
In this model, the pre-authentication signaling path between a peer and a candidate authenticator consists of two segments: peer-to-SAP signaling (over L2 or L3) and SAP-to-CAP signaling over L3.

### 6.2. The Authenticated Anticipatory Keying Usage Model

In this model, it is assumed that there is no trust relationship between the SAP and the CAP, and the SAP is required to interact with the AAA server directly. The authenticated anticipatory keying usage model is illustrated in Figure 6.

![Figure 6: Authenticated Anticipatory Keying Usage Model](image)

The role of the serving attachment point in this usage model is to communicate with the peer on one side and exchange authenticated anticipatory keying signaling with the EAP server on the other side. The role of the candidate authenticator is to receive the transported keying materials from the EAP server and to act as the serving attachment point after handover occurs. The MD-SAP signaling is performed over L2 or L3; the SAP-AAA and AAA-CAP segments operate over L3.

### 7. Architectural Considerations

There are two architectural issues relating to early authentication: authenticator discovery and context binding.

#### 7.1. Authenticator Discovery

In general, early authentication requires the identity of a candidate attachment point to be discovered by a peer, by a serving attachment point, or by some other entity prior to handover. An attachment point discovery protocol is typically defined as a separate protocol.
from an early authentication protocol. For example, the IEEE 802.21 Information Service (IS) [IEEE.802-21] provides a link-layer-independent mechanism for obtaining neighboring network information by defining a set of Information Elements (IEs), where one of the IEs is defined to contain an IP address of an attachment point. IEEE 802.21 IS queries for such an IE may be used as a method for authenticator discovery.

If IEEE 802.21 IS or a similar mechanism is used, authenticator discovery requires a database of information regarding the target network; the provisioning of a server with such a database is another issue.

7.2. Context Binding

When a candidate authenticator uses different EAP transport protocols for normal authentication and early authentication, a mechanism is needed to bind link-layer-independent context carried over early authentication signaling to the link-layer-specific context of the link to be established between the peer and the candidate authenticator. The link-layer-independent context includes the identities of the peer and authenticator as well as the MSK. The link-layer-specific context includes link-layer addresses of the peer and the candidate authenticator. Such context binding can happen before or after the peer changes its point of attachment.

There are at least two possible approaches to address the context binding issue. The first approach is based on communicating the link-layer context as opaque data via early authentication signaling. The second approach is based on running EAP over the link layer of the candidate authenticator after the peer arrives at the authenticator, using short-term credentials generated via early authentication. In this case, the short-term credentials are shared between the peer and the candidate authenticator. In both approaches, context binding needs to be securely made between the peer and the candidate authenticator. Also, the peer is not fully authorized by the candidate authenticator until the peer completes the link-layer-specific secure association procedure with the authenticator using link-layer signaling.

8. AAA Issues

Most of the AAA documents today do not distinguish between a normal authentication and an early authentication, and this creates a set of open issues:
Early authentication authorization

Users may not be allowed to have more than one logon session at the time. This means that while such users actively engage in a session (as a result of a previously valid authentication), they will not be able to perform early authentication. The AAA server currently has no way of distinguishing between a normal authentication request and an early authentication request.

Early authentication lifetime

Currently, AAA protocols define attributes carrying lifetime information for a normal authentication session. Even when a user profile and the AAA server support early authentication, the lifetime for an early authentication session is typically valid only for a short amount of time because the peer has not completed its authentication at the target link layer. It is currently not possible for a AAA server to indicate to the AAA client or a peer the lifetime of the early authenticated session unless AAA protocols are extended to carry early authentication session lifetime information. In other words, it is not clear to the peer or the authenticator when the early authentication session will expire.

Early authentication retries

It is typically expected that, shortly following the early authentication process, the peer moves to the new point of attachment and converts the early authentication state to a normal authentication state (the procedure for which is not the topic of this particular subsection). However, if the peer has not yet moved to the new location and realizes that the early authentication session is expiring, it may perform another early authentication. Some limiting mechanism is needed to avoid an unlimited number of early authentication attempts.

Completion of network attachment

Once the peer has successfully attached to the new point of attachment, it needs to convert its authentication state from early authenticated to fully attached and authorized. If the AAA server needs to differentiate between early authentication and normal authentication, there may need to be a mechanism within the AAA protocol to provide this indication to the AAA server. This may be important from a billing perspective if the billing policy does not charge for an early authenticated peer until the peer is fully attached to the target authenticator.

Session resumption

In the case where the peer cycles between a network N1 with which it has fully authenticated and another network N2 and then back to N1, it should be possible to simply convert the fully...
authenticated state on N1 to an early authenticated state. The problems around handling session lifetime and keying material caching need to be dealt with.

Multiple candidate attachment points
There may be situations where the peer needs to choose from a number of CAPs. In such cases, it is desirable for the peer to perform early authentication with multiple candidate authenticators. This amplifies the difficulties noted under the point "Early authentication authorization".

Inter-AAA-realm handover support
There may be situations where the peer moves out of the home AAA realm or across different visited AAA realms. In such cases, the early authentication should be performed through the visited AAA realm with the AAA server in the home AAA realm. It also requires AAA in the visited realm to acquire the identity information of the home AAA realms for routing the EAP early authentication traffic. Knowledge of realm identities is required by both the peer and AAA to generate the early authentication key for mutual authentication between the peer and the visited AAA server.

Inter-technology support
Current specifications on early authentication mostly deal with homogeneous 802.11 networks. AAA attributes such as Calling-Station-ID [RADEXT-WLAN] may need to be expanded to cover other access technologies. Furthermore, inter-technology handovers may require a change of the peer identifier as part of the handover. Investigation on the best type of identifiers for peers that support multiple access technologies is required.

9. Security Considerations

This section specifically covers threats introduced to the EAP model by early authentication. Security issues on general EAP and handover are described in other documents such as [RFC3748], [RFC4962], [RFC5169], and [RFC5247].

Since early authentication, as described in this document, needs to work across multiple attachment points, any solution needs to consider the following security threats.

First, a resource consumption denial-of-service attack is possible, where an attacker that is not on the same IP link as the legitimate peer or the candidate authenticator may send unprotected early authentication messages to the legitimate peer or the candidate authenticator. As a result, the latter may spend computational and bandwidth resources on processing early authentication messages sent
by the attacker. This attack is possible in both the direct and indirect pre-authentication scenarios. To mitigate this attack, the candidate network or authenticator may apply non-cryptographic packet filtering so that only early authentication messages received from a specific set of serving networks or authenticators are processed. In addition, a simple solution for the peer side would be to let the peer always initiate EAP early authentication and not allow EAP early authentication initiation from an authenticator.

Second, consideration for the channel binding problem described in [RFC5247] is needed as lack of channel binding may enable an authenticator to impersonate another authenticator or communicate incorrect information via out-of-band mechanisms (such as via a AAA or lower-layer protocol) [RFC3748]. It should be noted that it is relatively easier to launch such an impersonation attack for early authentication than normal authentication because an attacker does not need to be physically on the same link as the legitimate peer to send an early authentication trigger to the peer.

10. Acknowledgments

The editors would like to thank Preetida Vinayakray, Shubhranshu Singh, Ajay Rajkumar, Rafa Marin Lopez, Jong-Hyouk Lee, Maryna Komarova, Katrin Hoeper, Subir Das, Charles Clancy, Jari Arkko, and Bernard Aboba for their valuable input.

11. Contributors

The following people all contributed to this document: Alper E. Yegin, Tom Taylor, Srinivas Sreemanthula, Madjid Nakhjiri, Mahalingam Mani, and Ashutosh Dutta.

12. References

12.1. Normative References


12.2. Informative References


Authors’ Addresses

Yoshihiro Ohba
Toshiba Corporate Research and Development Center
1 Komukai-Toshiba-cho
Saiwai-ku, Kawasaki, Kanagawa, 212-8582
Japan
Phone: +81 44 549 2230
EMail: yoshihiro.ohba@toshiba.co.jp

Qin Wu (editor)
Huawei Technologies Co., Ltd
Huawei Nanjing R&D Center, Floor 1F, Software Avenue, No.101.,
Yuhua District
Nanjing, JiangSu 210012
China
Phone: +86 25 56622908
EMail: sunseawq@huawei.com

Glen Zorn (editor)
Network Zen
1463 East Republican Street
Seattle, Washington 98112
USA
EMail: gwz@net-zen.net