MOBOPTS Research Group
Internet-Draft
Intended status: Informational
Expires: August 26, 2011

A. Dutta (Ed.)
V. Fajardo
Telcordia
Y. Ohba
K. Taniuchi
Toshiba
H. Schulzrinne
Columbia Univ.
February 22, 2011

A Framework of Media-Independent Pre-Authentication (MPA) for Inter-domain Handover Optimization
draft-irtf-mobopts-mpa-framework-09

Abstract

This document describes Media-independent Pre-Authentication (MPA), a new handover optimization mechanism that addresses the issues on existing mobility management protocols and mobility optimization mechanisms to support inter-domain handover. MPA is a mobile-assisted, secure handover optimization scheme that works over any link-layer and with any mobility management protocol and is best applicable to support optimization during inter-domain handover. MPA's pre-authentication, pre-configuration, and proactive handover techniques allow many of the handoff related operations to take place before the mobile has moved to the new network. We describe the details of all the associated techniques and its applicability for different scenarios involving various mobility protocols during inter-domain handover. We have implemented MPA mechanism for various network layer and application layer mobility protocols and report summary of experimental performance results in this document.

This document is a product of the IP Mobility Optimizations (MobOpts) Research Group.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any

time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 26, 2011.

Copyright Notice

Table of Contents

1. Introduction

As wireless technologies including cellular and wireless LAN are beginning to get popular, supporting terminal handovers across different types of access networks, such as from a wireless LAN to CDMA or to GPRS is considered a clear challenge. On the other hand, supporting seamless terminal handovers between access networks of the same type is still more challenging, especially when the handovers are across IP subnets or administrative domains. To address those challenges, it is important to provide terminal mobility that is agnostic to link-layer technologies in an optimized and secure fashion without incurring unreasonable complexity. In this document we discuss a framework to support terminal mobility that provides seamless handovers with low latency and low loss. Seamless handovers are characterized in terms of performance requirements as described in Section 1.2. [mpa-wireless] is an accompanying document which describes implementation of a few MPA-based systems including performance results to show how existing protocols could be leveraged to realize the functionalities of MPA.

Terminal mobility is accomplished by a mobility management protocol that maintains a binding between a locator and an identifier of a mobile node, where the binding is referred to as the mobility binding. The locator of the mobile node may dynamically change when there is a movement of the mobile node. The movement that causes a change of the locator may occur when there is a change in attachment point due to physical movement or network change. A mobility management protocol may be defined at any layer. In the rest of this document, the term "mobility management protocol" refers to a mobility management protocol which operates at the network layer or higher.

There are several mobility management protocols at different layers. Mobile IP [RFC5944] and Mobile IPv6 [RFC3775] are mobility management protocols that operate at the network layer. Similarly, MOBIKE (IKEv2 Mobility and Multihoming) [RFC4555] is an extension to IKEv2 that provides the ability to deal with a change of an IP address of an IKEv2 end-point. There are several ongoing activities in the IETF to define mobility management protocols at layers higher than network layer. HIP (the Host Identity Protocol) [RFC5201] defines a new protocol layer between network layer and transport layer to provide terminal mobility in a way that is transparent to both network layer and transport layer. Also, SIP-based mobility is an extension to SIP to maintain the mobility binding of a SIP user agent [SIPMM].

While mobility management protocols maintain mobility bindings, these cannot provide seamless handover if used in their current form. An additional optimization mechanism is needed to prevent the loss of

inflight packets transmitted during mobile's binding update procedure and to achieve seamless handovers. Such a mechanism is referred to as a mobility optimization mechanism. For example, mobility optimization mechanisms [RFC4881] and [RFC5568] are defined for Mobile IPv4 and Mobile IPv6, respectively, by allowing neighboring access routers to communicate and carry information about mobile terminals. There are protocols that are considered as "helpers" of mobility optimization mechanisms. The CARD (Candidate Access Router Discovery Mechanism) protocol [RFC4065] is designed to discover neighboring access routers. The CTP (Context Transfer Protocol) [RFC4066] is designed to carry state that is associated with the services provided for the mobile node, or context, among access routers. We describe some of the fast-handover schemes that attempt to reduce the handover delay in Section 4.

There are several issues in existing mobility optimization mechanisms. First, existing mobility optimization mechanisms are tightly coupled with specific mobility management protocols. For example, it is not possible to use mobility optimization mechanisms designed for Mobile IPv4 or Mobile IPv6 for MOBIKE. What is strongly desired is a single, unified mobility optimization mechanism that works with any mobility management protocol. Second, there is no existing mobility optimization mechanism that easily supports handovers across administrative domains without assuming a pre-established security association between administrative domains. A mobility optimization mechanism should work across administrative domains in a secure manner only based on a trust relationship between a mobile node and each administrative domain. Third, a mobility optimization mechanism needs to support not only terminals with multiple-interfaces where simultaneous connectivity through multiple interfaces or connectivity through single interface can be expected, but also terminals with single-interface.

This document describes a framework of Media-independent Pre-Authentication (MPA), a new handover optimization mechanism that addresses all those issues. MPA is a mobile-assisted, secure handover optimization scheme that works over any link-layer and with any mobility management protocol including Mobile IPv4, Mobile IPv6, MOBIKE, HIP, SIP mobility. In cases of multiple operators without roaming relationship or without agreement to participate in a key management scheme, MPA provides a framework that can perform pre-authentication to establish the security mechanisms without assuming a common source of trust. In MPA, the notion of IEEE 802.11i pre-authentication is extended to work at higher layer, with additional mechanisms to perform early acquisition of IP address from a network where the mobile node may move as well as proactive handover to the network while the mobile node is still attached to the current network. Since this document focuses on the MPA framework, it is

left to future work to choose the protocols for MPA and define detailed operations. The accompanying document [mpa-wireless] provides one method that describes usage and interactions between existing protocols to accomplish MPA functionality.

This document represents the consensus of the (MobOpts) Research Group. It has been reviewed by Research Group members active in the specific area of work.

1.1. Specification of Requirements

In this document, several words are used to signify the requirements of the specification. These words are often capitalized. The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

1.2. Performance Requirements

In order to provide desirable quality of service for interactive VoIP and streaming traffic, one needs to limit the value of end-to-end delay, jitter and packet loss to a certain threshold level. ITU-T and ITU-E standards define the acceptable values for these parameters. For example for one-way delay, ITU-T G.114 [RG98] recommends 150 ms as the upper limit for most of the applications, and 400 ms as generally unacceptable delay. One way delay tolerance for video conferencing is in the range of 200 to 300 ms [ITU98]. Also if an out-of-order packet is received after a certain threshold, it is considered lost. According to ETSI TR 101 [ETSI], a normal voice conversation can tolerate up to 2% packet loss. But this is the mean packet loss probability and may be applicable to a scenario when the mobile is subjected to repeated handoff during a normal conversation. Measurement techniques for delay and jitter are described in [RFC2679], [RFC2680] and [RFC2681].

In case of interactive VoIP traffic, end-to-end delay affects the jitter value, and thus is an important issue to consider. An end-to-end delay consists of several components, such as network delay, operating system (OS) delay, codec delay and application delay. A complete analysis of these delays can be found in [Wenyu]. During a mobile's handover, in-flight transient traffic cannot reach the mobile because of the associated handover delay. These in-flight packets could either be lost or buffered. If the in-flight packets are lost, then it contributes to jitter between the last packet before handoff and first packet after handoff. If these packets are buffered, packet loss is minimized, but there is additional jitter for the in-flight packets when these are flushed after the handoff. Buffering during handoff avoids the packet loss, but at the cost of

additional one-way-delay. A trade-off between one-way-delay and packet loss is desired based on the type of application. For example, for streaming application, packet loss can be reduced by increasing the playout buffer resulting in longer one-way packet delay.

The handover delay is attributed due to several factors, such as discovery, configuration, authentication, binding update and media delivery. Many of the security related procedures such as handover keying and re-authentication procedures deal with cases where there is a single source of trust at the top and the underlying AAA domain elements trust the top source of trust and the keys it generates and distributes. In this scenario, there is an appreciable delay in re-establishing link security related parameters, such as authentication, link key management and access authorization during inter-domain handover. The focus of this draft is the design of a framework that can reduce the delay due to authentication and other handoff related operations such as configuration and binding update.

2. Terminology

Mobility Binding: A binding between a locator and an identifier of a mobile terminal.

Mobility Management Protocol (MMP): A protocol that operates at network layer or above to maintain a binding between a locator and an identifier of a mobile node.

Binding Update: A procedure to update a mobility binding.

Media-independent Pre-Authentication Mobile Node (MN): A mobile node of media-independent pre-authentication (MPA) which is a mobile-assisted, secure handover optimization scheme that works over any link-layer and with any mobility management protocol. An MPA mobile node is an IP node. In this document, the term "mobile node" or "MN" without a modifier refers to "MPA mobile node". An MPA mobile node usually has a functionality of a mobile node of a mobility management protocol as well.

Candidate Target Network (CTN):

A network to which the mobile may move in the near future.

Target Network (TN): The network to which the mobile has decided to move. The target network is selected from one or more candidate target network.

Proactive Handover Tunnel (PHT): A bidirectional IP tunnel [RFC2003], [RFC2473] that is established between the MPA mobile node and an access router of a candidate target network. In this document, the term "tunnel" without a modifier refers to "proactive handover tunnel.

Point of Attachment (PoA): A link-layer device (e.g., a switch, an access point or a base station) that functions as a link-layer attachment point for the MPA mobile node to a network.

Care-of Address (CoA): An IP address used by a mobility management protocol as a locator of the MPA mobile node.

3. Handover Taxonomy

Based on the type of movement, type of access network, and underlying mobility support, one can primarily define the handover as inter-technology, intra-technology, inter-domain, and intra-domain. We describe briefly each of these handover processes. However, our focus of the dicussion is on Inter-domain handover.

Inter-technology: A mobile may be equipped with multiple interfaces, where each interface can support different access technology (802.11, CDMA). A mobile may communicate with one interface at any time in order to conserve the power. During the handover the mobile may move out of the footprint of one access technology (e.g., 802.11) and move into the footprint of a different access technology (e.g., CDMA). This will warrant switching of the communicating interface on the mobile as well. This type of Inter-technology handover is often called as Vertical Handover since the mobile makes movement between two different cell sizes.

Intra-technology: An intra-technology handover is defined when a mobile moves between the same type of access technology such as between 802.11[a,b,n] and 802.11 [a,b,n] or between CDMA1XRTT and CDMA1EVDO. In this scenario a mobile may be equipped with a single interface (with multiple PHY types of the same technology) or with multiple interfaces. An Intra-technology handover may involve intra-subnet or inter-subnet movement and thus may need to change its L3 locator depending upon type of movement.

Inter-domain: A domain can be defined in several ways. But for the purposes of roaming we define domain as an administrative domain

which consists of networks that are managed by a single administrative entity which authenticates and authorizes a mobile for accessing the networks. An administrative entity may be a service provider, an enterprise and any organization. Thus an Inter-domain handover will by-default be subjected to inter-subnet handover and in addition it may be subjected to either inter-technology or intra-technology handover. A mobile is subjected to inter-subnet handover when it moves from one subnet (broadcast domain) to another subnet (broadcast domain). Inter-domain handover will be subjected to all the transition steps a subnet handover goes through and in addition it will be subjected to authentication and authorization process as well. It is also likely that the type of mobility support in each administrative domain will be different. For example, administrative domain A may have MIPv6 support, while administrative domain B may use Proxy MIPv6.

Intra-domain: When a mobile's movement is confined to movement within an administrative domain it is called intra-domain movement. An intra-domain movement may involve intra-subnet, inter-subnet, intra-technology and inter-technology as well.

Both inter-domain and intra-domain handovers can be subjected to either inter-technology or intra-technology handover based on the network access characteristics. Inter-domain handover requires authorization for acquisition or modification of resources assigned to a mobile and the authorization needs interaction with a central authority in a domain. In many cases, an authorization procedure during inter-domain handover follows an authentication procedure that also requires interaction with a central authority in a domain. Thus, security associations between the network entities such as routers in the neighboring administrative domains need to be established before any interaction takes place between these entities. Similarly, an inter-domain mobility may involve different mobility protocols in each of its domains, such as MIPv6 and Proxy-MIPv6. In that case, one needs a generalized framework to achieve the optimization during inter-domain handover. Figure 1 shows a typical example of inter-domain mobility involving two domains, such as domain A and domain B. It illustrates several important components such as AAA Home server (AAAH), AAA visited servers (e.g., AAAV1 and AAAV2), Authentication Agent (AA), Layer 3 point of attachment, such as Access Router (AR) and layer 2 point of attachment, such as Access Point. Any mobile maybe using a specific mobility protocol and associated mobility optimization technique during intra-domain movement in either domain. But the same optimization technique may not be suitable to support inter-domain handover independent of whether it uses the same or different mobility protocol in either domain.

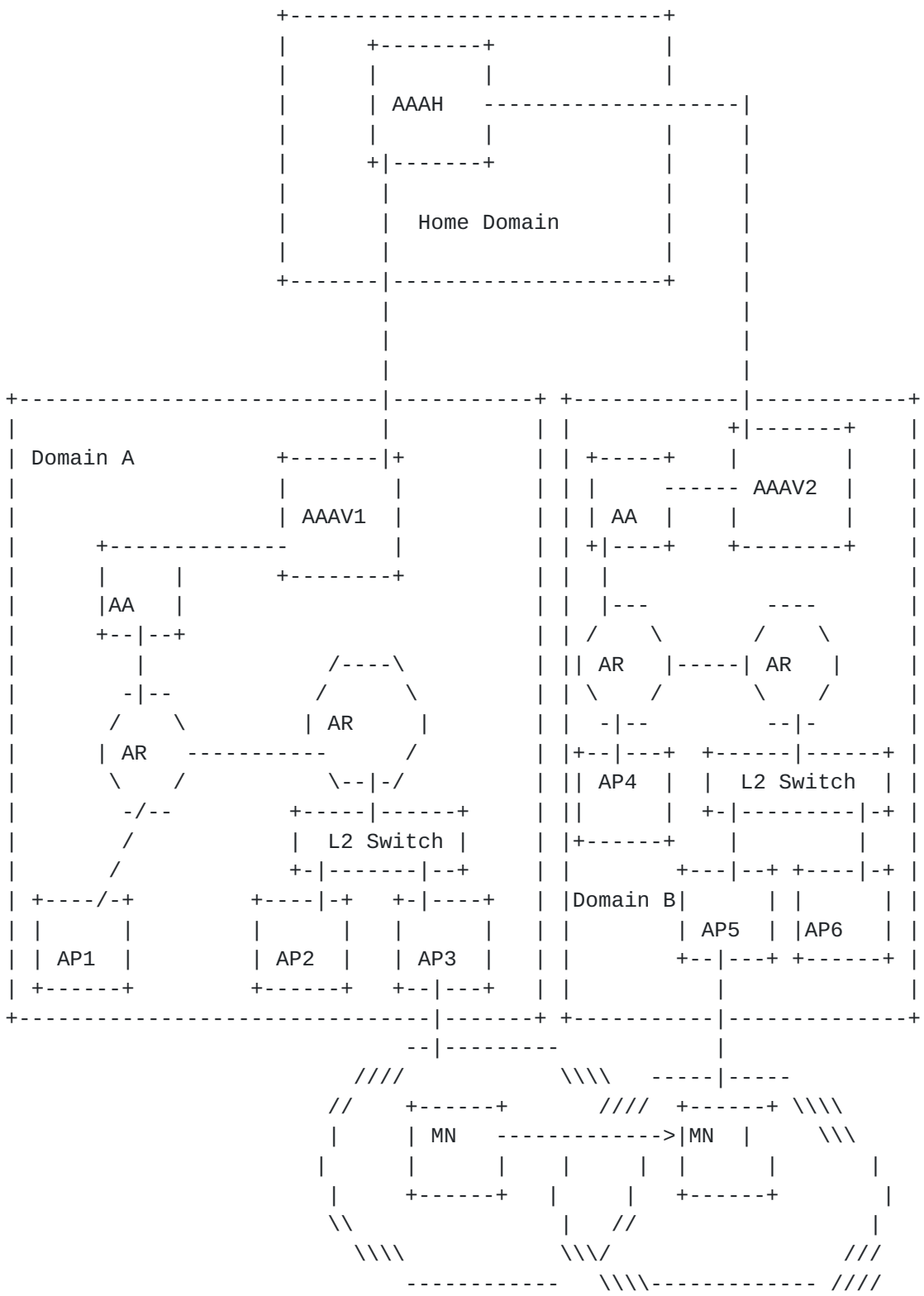

Figure 1: Inter-domain Mobility

4. Related Work

While basic mobility management protocols such as Mobile IP [RFC5944], Mobile IPv6 [RFC3775], SIP-Mobility [SIPMM] provide continuity to TCP and RTP traffic, these are not optimized to reduce the handover latency during mobile's movement between subnets and domains. In general these mobility management protocols introduce handover delays incurred at several layers such as, layer 3 and application layer for updating the mobile's mobility binding. These protocols also get affected due to underlying layer 2 delay as well. As a result, applications using these mobility protocols suffer from performance degradation.

There have been several optimization techniques that apply to current mobility management schemes that try to reduce handover delay and packet loss during a mobile's movement between cells, subnets and domain. Micro-mobility management schemes [CELLIP], [HAWAII], and intra-domain mobility management schemes such as [IDMP], [I-D.ietf-mobileip-reg-tunnel] and [RFC5380] provide fast-handover by limiting the signaling updates within a domain. Fast Mobile IP protocols for IPv4 and IPv6 networks [RFC4881], [RFC5568] utilize mobility information made available by link layer triggers. Yokota et al. [YOKOTA] propose joint use of access point and a dedicated MAC bridge to provide fast-handover without altering the MIPv4 specification. Shin et al. [MACD] propose a scheme reducing the delay due to MAC layer handoff by providing a cache-based algorithm. In this scheme, the mobile caches the neighboring channels that it has already visited and thus uses a selective scanning method. This helps to reduce the associated scanning time.

Some mobility management schemes use dual interfaces thus providing make-before-break [SUM]. In a make-before-break situation, communication usually continues with one interface, when the secondary interface is in the process of getting connected. The IEEE 802.21 working group is discussing these scenarios in detail [802.21]. Providing fast-handover using a single interface needs more careful design than for a client with multiple interfaces. Dutta et al [SIPFAST] provide an optimized handover scheme for SIP-based mobility management, where the transient traffic is forwarded from the old subnet to the new one by using an application layer forwarding scheme. [MITH] provides a fast handover scheme for the single interface case that uses mobile-initiated tunneling between the old foreign agent and new foreign agent. [MITH] defines two types of handover schemes such as Pre-MIT (Mobile Initiated Tunneling) and Post-MIT (Media Initiated Tunneling). The proposed MPA scheme is very similar to MITH's predictive scheme where the mobile communicates with the foreign agent before actually moving to the new network. However, the MPA scheme is not limited to MIP; this

scheme takes care of movement between domains and performs pre-authentication in addition to proactive handover. Thus, MPA reduces the overall delay to close to link-layer handover delay. Most of the mobility optimization techniques developed so far are restricted to a specific type of mobility protocol only. While supporting optimization for inter-domain mobility, these protocols assume that there is a pre-established security arrangement between two administrative domains. But this assumption may not be viable always. Thus, there is a need to develop an optimization mechanism that can support inter-domain mobility without any underlying constraints or security related assumption.

Recently, the HOKEY WG within IETF is defining the ways to expedite the authentication process. In particular, it has defined pre-authentication [RFC5836] and fast re-authentication [RFC5169] mechanisms to expedite the authentication and security association process.

5. Applicability of MPA

MPA is more applicable where an accurate prediction of movement can be easily made. For other environments, special care must be taken to deal with issues such as pre-authentication to multiple CTNs (Candidate Target Networks) and failed switching and switching back as described in [mpa-wireless]. However, addressing those issues in actual deployments may not be easier. Some of the deployment issues are described in Appendix C.

Authors have cited several use cases of how MPA can be used to optimize several network layer and application layer mobility protocols in an accompanying document [mpa-wireless]. The effectiveness of MPA may be relatively reduced if the network employs network-controlled localized mobility management in which the MN does not need to change its IP address while moving within the network. The effectiveness of MPA may also be relatively reduced if signaling for network access authentication is already optimized for movements within the network, e.g., when simultaneous use of multiple interfaces during handover is allowed. In other words, MPA is a more viable as a solution for inter-administrative domain predictive handover without the simultaneous use of multiple interfaces. Since MPA is not tied to a specific mobility protocol, it is also applicable to support optimization for inter-domain handover where each domain may be equipped with a different mobility protocol. Figure 1 shows an example of inter-domain mobility where MPA could be applied. For example, domain A may support just Proxy MIPv6, whereas domain B may support Client Mobile IPv6. MPA's different functional components can provide the desired optimization techniques

proactively.

6. MPA Framework

6.1. Overview

Media-independent Pre-Authentication (MPA) is a mobile-assisted, secure handover optimization scheme that works over any link layer and with any mobility management protocol. With MPA, a mobile node is not only able to securely obtain an IP address and other configuration parameters for a CTN, but also able to send and receive IP packets using the IP address obtained before it actually attaches to the CTN. This makes it possible for the mobile node to complete the binding update of any mobility management protocol and use the new CoA before performing a handover at link-layer.

MPA adopts the following basic procedures to provide this functionality. The first procedure is referred to as "pre-authentication", the second procedure is referred to as "pre-configuration", the combination of the third and fourth procedures are referred to as "secure proactive handover". The security association established through pre-authentication is referred to as an "MPA-SA".

This functionality is provided by allowing a mobile node which has connectivity to the current network but is not yet attached to a CTN, to (i) establish a security association with the CTN to secure the subsequent protocol signaling, then (ii) securely execute a configuration protocol to obtain an IP address and other parameters from the CTN as well as execute a tunnel management protocol to establish a Proactive Handover Tunnel (PHT) [RFC2003] between the mobile node and an access router of the CTN, then (iii) send and receive IP packets, including signaling messages for binding update of an MMP and data packets transmitted after completion of binding update, over the PHT using the obtained IP address as the tunnel inner address, and finally (iv) deleting or disabling the PHT immediately before attaching to the CTN when it becomes the target network and then re-assigning the inner address of the deleted or disabled tunnel to its physical interface immediately after the mobile node is attached to the target network through the interface. Instead of deleting or disabling the tunnel before attaching to the target network, the tunnel may be deleted or disabled immediately after being attached to the target network.

Especially, the step (iii) in the previous paragraph (i.e., binding update procedure) makes it possible for the mobile to complete the higher-layer handover before starting link-layer handover. This

means that the mobile is able to send and receive data packets transmitted after completing the binding update over the tunnel, while data packets transmitted before completion of binding update do not use the tunnel.

6.2. Functional Elements

In the MPA framework, the following functional elements are expected to reside in each CTN to communicate with a mobile node: Authentication Agent (AA), Configuration Agent (CA) and Access Router (AR). These elements can reside in one or more network devices.

An authentication agent is responsible for pre-authentication. An authentication protocol is executed between the mobile node and the authentication agent to establish an MPA-SA. The authentication protocol MUST be able to establish a shared key between the mobile node and the authentication agent and SHOULD be able to provide mutual authentication. The authentication protocol SHOULD be able to interact with a AAA protocol such as RADIUS and Diameter to carry authentication credentials to an appropriate authentication server in the AAA infrastructure. This interaction happens through the Authentication Agent such as PANA Authentication Agent (PAA). The derived key is used for further deriving keys used for protecting message exchanges used for pre-configuration and secure proactive handover. Other keys that are used for bootstrapping link-layer and/or network-layer ciphers MAY also be derived from the MPA-SA. A protocol that can carry EAP [RFC3748] would be suitable as an authentication protocol for MPA.

A configuration agent is responsible for one part of pre-configuration, namely securely executing a configuration protocol to deliver an IP address and other configuration parameters to the mobile node. The signaling messages of the configuration protocol (e.g., DHCP) MUST be protected using a key derived from the key corresponding to the MPA-SA.

An access router in MPA framework is a router that is responsible for the other part of pre-configuration, i.e., securely executing a tunnel management protocol to establish a proactive handover tunnel to the mobile node. IP packets transmitted over the proactive handover tunnel SHOULD be protected using a key derived from the key corresponding to the MPA-SA. Details of this procedure are described in Section 6.3.

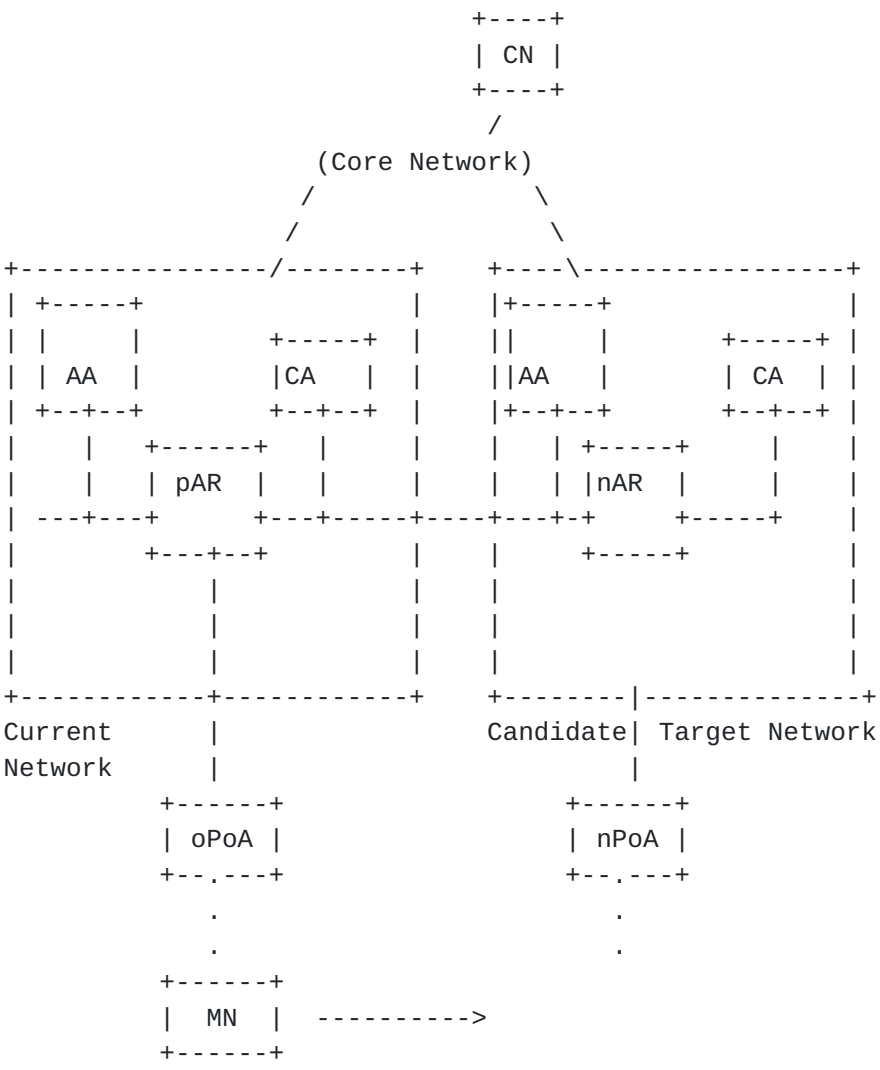

Figure 2: MPA Functional Components

6.3. Basic Communication Flow

Assume that the mobile node is already connected to a point of attachment, say oPoA (old point of attachment), and assigned a care-of address, say oCoA (old care-of address). The communication flow of MPA is described as follows. Throughout the communication flow, data packet loss should not occur except for the period during the switching procedure in Step 5, and it is the responsibility of link-layer handover to minimize packet loss during this period.

Step 1 (pre-authentication phase): The mobile node finds a CTN through some discovery process such as IEEE 802.21 and obtains the IP

addresses of an authentication agent, a configuration agent and an access router in the CTN (Candidate Target Network) by some means. Details about discovery mechanisms are discussed in Section 7.1. The mobile node performs pre-authentication with the authentication agent. As discussed in Section 7.2, the mobile may need to pre-authenticate with multiple candidate target networks. The decision regarding which candidate network the mobile needs to pre-authenticate with will depend upon several factors, such as signaling overhead, bandwidth requirement (QoS), mobile's location, communication cost, and handover robustness etc. Determining the policy that decides the target network the mobile should pre-authenticate with is out of scope for this document.

If the pre-authentication is successful, an MPA-SA is created between the mobile node and the authentication agent. Two keys are derived from the MPA-SA, namely an MN-CA key and an MN-AR key, which are used to protect subsequent signaling messages of a configuration protocol and a tunnel management protocol, respectively. The MN-CA key and the MN-AR key are then securely delivered to the configuration agent and the access router, respectively.

Step 2 (pre-configuration phase): The mobile node realizes that its point of attachment is likely to change from oPoA to a new one, say nPoA (new point of attachment). It then performs pre-configuration with the configuration agent using the configuration protocol to obtain several configuration parameters such as an IP address, say nCoA (new care-of address), and default router from the CTN. The mobile then communicates with the access router using the tunnel management protocol to establish a proactive handover tunnel. In the tunnel management protocol, the mobile node registers oCoA and nCoA as the tunnel outer address and the tunnel inner address, respectively. The signaling messages of the pre-configuration protocol are protected using the MN-CA key and the MN-AR key. When the configuration and the access router are co-located in the same device, the two protocols may be integrated into a single protocol like IKEv2. After completion of the tunnel establishment, the mobile node is able to communicate using both oCoA and nCoA by the end of Step 4. A configuration protocol and a tunnel management protocol may be combined in a single protocol or executed in different orders depending on the actual protocol(s) used for configuration and tunnel management.

Step 3 (secure proactive handover main phase): The mobile node decides to switch to the new point of attachment by some means. Before the mobile node switches to the new point of attachment, it starts secure proactive handover by executing the binding update operation of a mobility management protocol and transmitting subsequent data traffic over the tunnel (main phase). This proactive

binding update could be triggered based on certain local policy at the mobile node end, after the pre-configuration phase is over. This local policy could be signal-to-noise ratio, location of the mobile etc. In some cases, it may cache multiple nCOA addresses and perform simultaneous binding with the CN or HA.

Step 4 (secure proactive handover pre-switching phase): The mobile node completes the binding update and becomes ready to switch to the new point of attachment. The mobile may execute the tunnel management protocol to delete or disable the proactive handover tunnel and cache nCoA after deletion or disabling of the tunnel. This transient tunnel can be deleted prior to or after the handover. The buffering module at the next access router buffers the packets once the tunnel interface is deleted. The decision as to when the mobile node is ready to switch to the new point of attachment depends on the handover policy.

Step 5 (switching): It is expected that a link-layer handover occurs in this step.

Step 6 (secure proactive handover post-switching phase): The mobile node executes the switching procedure. Upon successful completion of the switching procedure, the mobile node immediately restores the cached nCoA and assigns it to the physical interface attached to the new point of attachment. If the proactive handover tunnel was not deleted or disabled in Step 4, the tunnel is deleted or disabled as well. After this, direct transmission of data packets using nCoA is possible without using a proactive handover tunnel.

Call flow for MPA is shown in Figure 3 and Figure 4.

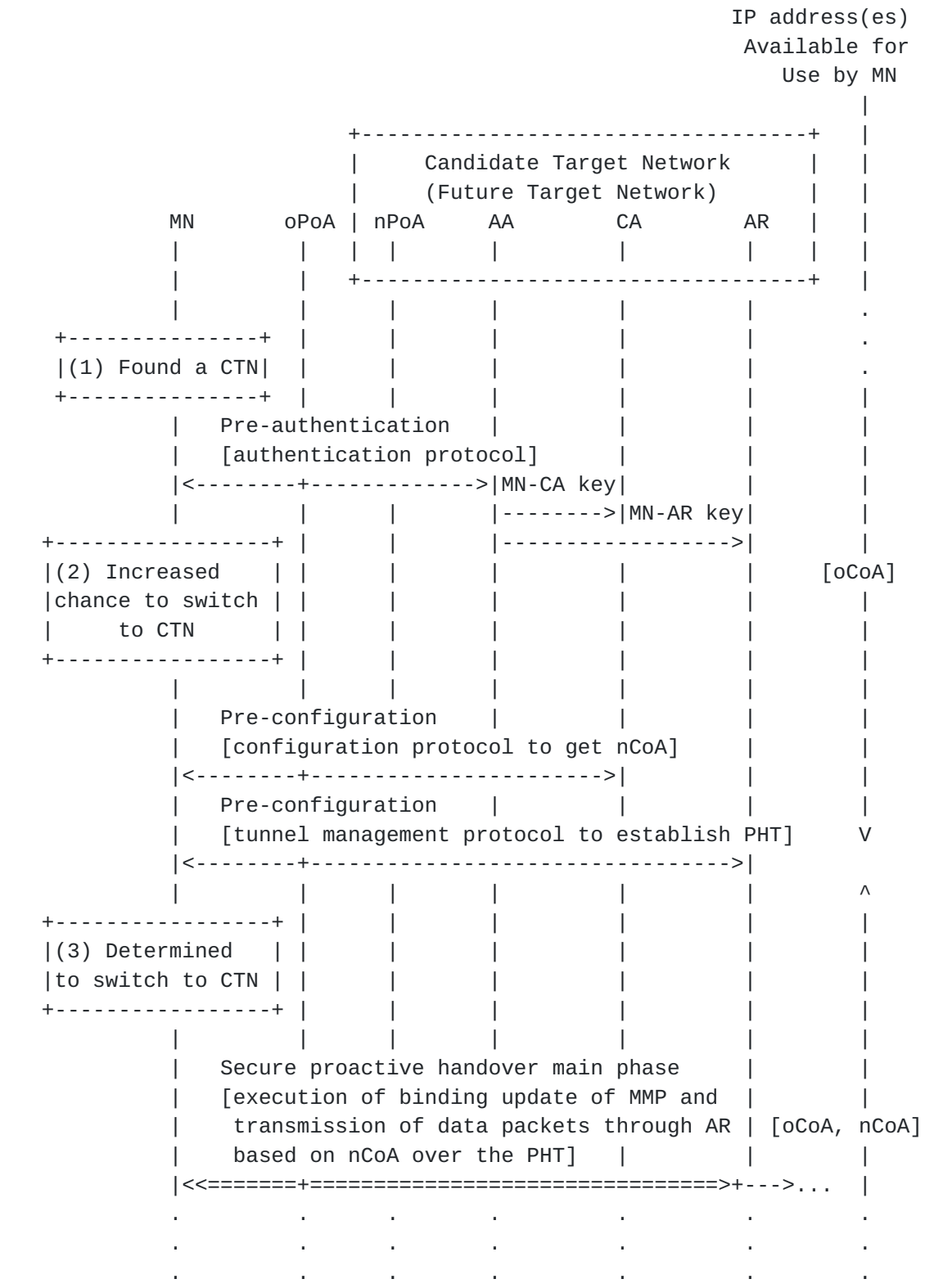

Figure 3: Example Communication Flow (1/2)

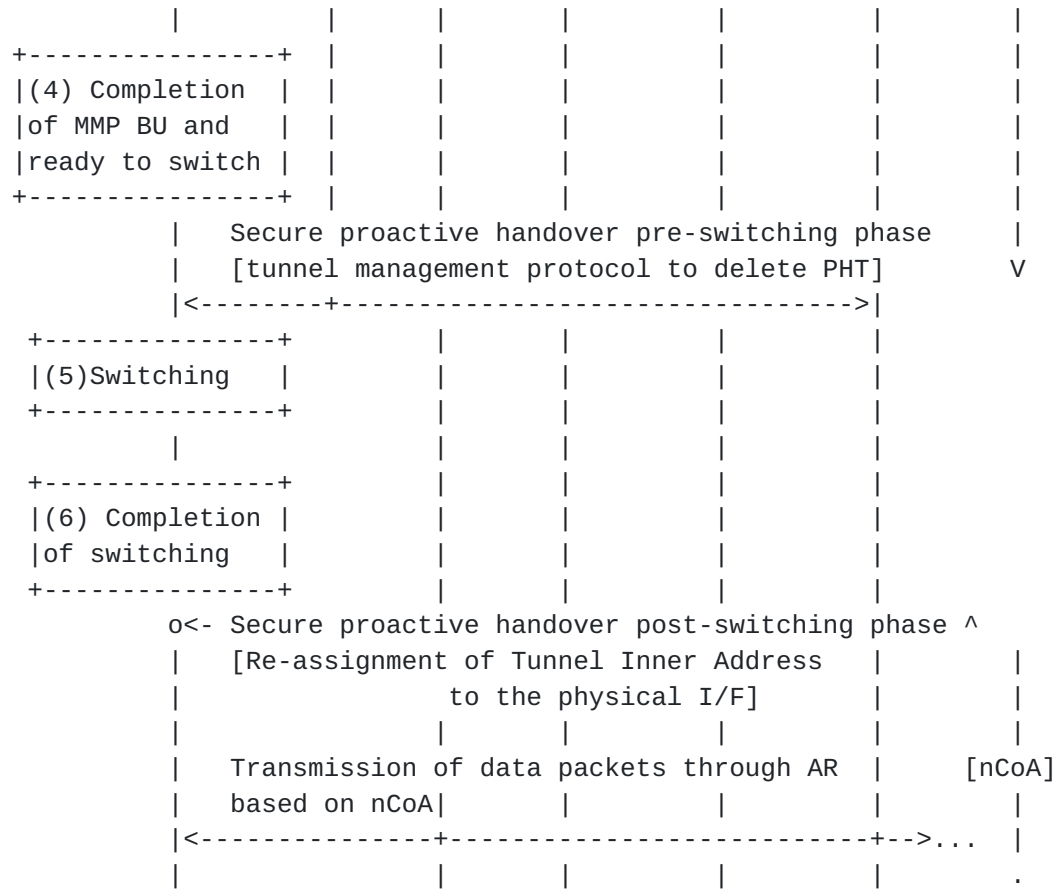

Figure 4: Example Communication Flow (2/2)

7. MPA Operations

In order to provide an optimized handover for a mobile experiencing rapid movement between subnets and/or domains handover, one needs to look into several operations. These issues include:

i) discovery of neighboring networking elements, ii) connecting to the right network based on certain policy, iii) changing the layer 2 point of attachment, iv) obtaining an IP address from a DHCP or PPP server, v) confirming the uniqueness of the IP address, vi) pre-authenticating with the authentication agent, vii) sending the binding update to the correspondent host viii) obtaining the redirected streaming traffic to the new point of attachment, ix) ping-pong effect, x) probability of moving to more than one network and associating with multiple target networks. We describe these issues in detail in the following paragraphs and describe how we have optimized these in case of MPA-based secure proactive handover.

7.1. Discovery

Discovery of neighboring networking elements such as access points, access routers, authentication servers helps expedite the handover process during a mobile's movement between networks. After discovering the network neighborhood with a desired set of coordinates, capabilities and parameters the mobile can perform many of the operation such as pre-authentication, proactive IP address acquisition, proactive address resolution, and binding update while in the previous network.

There are several ways a mobile can discover neighboring networks. The Candidate Access Router Discovery protocol [RFC2608] helps discover the candidate access routers in the neighboring networks. Given a certain network domain SLP (Service Location Protocol) [RFC4066] and DNS help provide addresses of the networking components for a given set of services in the specific domain. In some cases many of the network layer and upper layer parameters may be sent over link layer management frames such as beacons when the mobile approaches the vicinity of the neighboring networks. IEEE 802.11u is considering issues such as discovering neighborhood using information contained in link layer. However, if the link-layer management frames are encrypted by some link layer security mechanism, then the mobile node may not be able to obtain the requisite information before establishing link layer connectivity to the access point. In addition this may add burden to the bandwidth constrained wireless medium. In such cases a higher layer protocol is preferred to obtain the information regarding the neighboring elements. Some proposals such as [802.21] help obtain information about the neighboring networks from a mobility server. When the movement is imminent, the mobile node starts the discovery process by querying a specific server and obtains the required parameters such as the IP address of the access point, its characteristics, routers, SIP servers or authentication servers of the neighboring networks. In the event of multiple networks, it may obtain the required parameters from more than one neighboring networks and keep these in a cache. At some point the mobile finds out several CTNs out of many probable networks and starts the pre-authentication process by communicating with the required entities in the CTNs. Further details of this scenario are in Section 7.2.

7.2. Pre-authentication in multiple CTN environment

In some cases, although a mobile selects a specific network to be the target network, it may actually end up with moving into a neighboring network other than the target network due to factors that are beyond the mobile's control. Thus it may be useful to perform the pre-authentication with a few probable candidate target networks and

establish time-bound transient tunnels with the respective access routers in those networks. Thus, in the event of a mobile moving to a candidate target network other than that was chosen as the target network, it will not be subjected to packet loss due to authentication and IP address acquisition delay that could occur if the mobile did not pre-authenticate with that candidate target network. It may appear that by pre-authenticating with a number of candidate target networks and reserving the IP addresses, the mobile is reserving resources that could be used otherwise. But since this happens for a time-limited period it should not be a big problem; it depends upon the mobility pattern and duration. The mobile uses a pre-authentication procedure to obtain an IP address proactively and to set up the time bound tunnels with the access routers of the candidate target networks. Also, MN may retain some or all of the nCoAs for future movement.

The mobile may choose one of these addresses as the binding update address and send it to the CN (Correspondent Node) or HA (Home Agent), and will thus receive the tunneled traffic via the target network while in the previous network. But in some instances, the mobile may eventually end up moving to a network that is other than the target network. Thus, there will be a disruption in traffic as the mobile moves to the new network since the mobile has to go through the process of assigning the new IP address and sending the binding update again. There are two solutions to this problem. The mobile can take advantage of the simultaneous mobility binding and send multiple binding updates to the corresponding host or HA. Thus, the corresponding host or HA forwards the traffic to multiple IP addresses assigned to the virtual interfaces for a specific period of time. This binding update gets refreshed at the CH after the mobile moves to the new network, thus stopping the flow to the other candidate networks. RFC 5648 [RFC5648] discusses different scenarios of mobility binding with multiple care-of-addresses. In case simultaneous binding is not supported in a specific mobility scheme, forwarding of traffic from the previous target network will help take care of the transient traffic until the new binding update is sent from the new network.

7.3. Proactive IP address acquisition

In general a mobility management protocol works in conjunction with the Foreign Agent or in co-located address mode. The MPA approach can use both co-located address mode and foreign agent address mode. We discuss here the address assignment component that is used in co-located address mode. There are several ways a mobile node can obtain an IP address and configure itself. In some cases, a mobile can configure itself statically in the absence of any configuration element such as a server or router in the network. In a LAN

environment the mobile can obtain an IP address from DHCP servers. In case of IPv6 networks, a mobile has the option of obtaining the IP address using stateless auto-configuration or DHCPv6. In some wide area networking environment, the mobile uses PPP (Point-to-Point Protocol) to obtain the IP address by communicating with a NAS.

Each of these processes takes of the order of few hundred milliseconds to few seconds depending upon the type of IP address acquisition process and operating system of the clients and servers. Since IP address acquisition is part of the handover process, it adds to the handover delay and thus it is desirable to reduce this delay as much as possible. There are few optimized techniques such as DHCP Rapid Commit [RFC4039], GPS-coordinate based IP address [GPSIP] available that attempt to reduce the handover time due to IP address acquisition time. However, in all these cases the mobile also obtains the IP address after it moves to the new subnet and incurs some delay because of the signaling handshake between the mobile node and the DHCP server.

In FastMIP6 [RFC5568], through the RtSolPr and PrRtAdv messages, the MN also formulates a prospective new CoA (NCoA) when it is still present on the PAR's link. Hence, the latency due to new prefix discovery subsequent to handover is eliminated. However, in this case, both the previous access router (PAR) and the next access router (NAR) need to cooperate with each other to be able to retrieve the prefix from the target network.

In the following paragraph we describe few ways a mobile node can obtain the IP address proactively from the CTN and the associated tunnel setup procedure. These can broadly be divided into four categories such as PANA-assisted proactive IP address acquisition, IKE-assisted proactive IP address acquisition, proactive IP address acquisition using DHCP only and stateless autoconfiguration.When DHCP is used for address configuration, a DHCP server is assumed to be serving one subnet.

7.3.1. PANA-assisted proactive IP address acquisition

In case of PANA-assisted proactive IP address acquisition, the mobile node obtains an IP address proactively from a CTN. The mobile node makes use of PANA [RFC5191] messages to trigger the IP address acquisition process via a DHCP client that is colocated with the PANA authentication agent in the access router in the CTN acting on behalf of the mobile. Upon receiving a PANA message from the mobile node, the DHCP client on the authentication agent performs normal DHCP message exchanges to obtain the IP address from the DHCP server in the CTN. This address is piggy-backed in a PANA message and is delivered to the mobile. In case of IPv6 Router Advertisment (RA) is

carried as part of PANA message. In case of stateless autoconfiguration, the mobile uses the prefix(es) obtained as part of RA and its MAC address to construct the unique IPv6 address(es) as it would have done in the new network. In case of stateful address configuration, te procedure similar to DHCPv4 can be applied.

7.3.2. IKEv2-assisted proactive IP address acquisition

IKEv2-assisted proactive IP address acquisition works when an IPsec gateway and a DHCP relay agent are resident within each access router in the CTN. In this case, the IPsec gateway and DHCP relay agent in a CTN help the mobile node acquire the IP address from the DHCP server in the CTN. The MN-AR key established during the pre-authentication phase is used as the IKEv2 pre-shared secret needed to run IKEv2 between the mobile node and the access router. The IP address from the CTN is obtained as part of standard IKEv2 procedure, with using the co-located DHCP relay agent for obtaining the IP address from the DHCP server in the target network using standard DHCP. The obtained IP address is sent back to the client in the IKEv2 Configuration Payload exchange. In this case, IKEv2 is also used as the tunnel management protocol for a proactive handover tunnel (see Section 7.4). Alternatively VPN-GW can itself dispense the IP address from its IP address pool.

7.3.3. Proactive IP address acquisition using DHCPv4 only

As another alternative, DHCP may be used for proactively obtaining an IP address from a CTN without relying on PANA or IKEv2-based approaches by allowing direct DHCP communication between the mobile node and the DHCP relay or DHCP server in the CTN. The mechanism described in this section is applicable to DHCPv4 only. The mobile node sends a unicast DHCP message to the DHCP relay agent or DHCP server in the CTN requesting an address, while using the address associated with the current physical interface as the source address of the request.

When the message is sent to the DHCP relay agent, the DHCP relay agent relays the DHCP messages back and forth between the mobile node and the DHCP server. In the absence of a DHCP relay agent the mobile can also directly communicate with the DHCP server in the target network. The broadcast option in the client's unicast DISCOVER message should be set to 0 so that the relay agent or the DHCP server can send the reply directly back to the mobile using the mobile node's source address.

In order to prevent malicious nodes from obtaining an IP address from the DHCP server, DHCP authentication should be used or the access router should install a filter to block unicast DHCP message sent to

the remote DHCP server from mobile nodes that are not pre-authenticated. When DHCP authentication is used, the DHCP authentication key may be derived from the MPA-SA established between the mobile node and the authentication agent in the candidate target network.

The proactively obtained IP address is not assigned to the mobile node's physical interface until the mobile has moved to the new network. The IP address thus obtained proactively from the target network should not be assigned to the physical interface but rather to a virtual interface of the client. Thus, such a proactively acquired IP address via direct DHCP communication between the mobile node and the DHCP relay or the DHCP server in the CTN may be carried with additional information that is used to distinguish it from other addresses as assigned to the physical interface.

Upon the mobile's entry to the new network, the mobile node can perform DHCP over the physical interface to the new network to get other configuration parameters such as the SIP server, DNS server by using DHCP INFORM. This should not affect the ongoing communication between the mobile and correspondent host. Also, the mobile node can perform DHCP over the physical interface to the new network to extend the lease of the address that was proactively obtained before entering the new network.

In order to maintain the DHCP binding for the mobile node and keep track of the dispensed IP address before and after the secure proactive handover, the same DHCP client identifier needs to be used for the mobile node for both DHCP for proactive IP address acquisition and DHCP performed after the mobile node enters the target network. The DHCP client identifier may be the MAC address of the mobile node or some other identifier.

7.3.4. Proactive IP address acquisition using stateless autoconfiguration

For IPv6, a network address is configured either using DHCPv6 or stateless autoconfiguration. In order to obtain the new IP address proactively, the router advertisement of the next hop router can be sent over the established tunnel, and a new IPv6 address is generated based on the prefix and MAC address of the mobile. Generating a COA from the new network will avoid the time needed to obtain an IP address and perform Duplicate Address Detection.

Duplicate address detection and address resolution are part of the IP address acquisition process. As part of the proactive configuration these two processes can be done ahead of time. Details of how these two processes can be done proactively are described in Appendix A and

Appendix B, respectively.

In case of stateless autoconfiguration, the mobile checks to see the prefix of the router advertisement in the new network and matches it with the prefix of newly assigned IP address. If these turn out to be the same then the mobile does not go through the IP address acquisition phase again.

7.4. Tunnel management

After an IP address is proactively acquired from the DHCP server in a CTN or via stateless autoconfiguration in case of IPv6, a proactive handover tunnel is established between the mobile node and the access router in the CTN. The mobile node uses the acquired IP address as the tunnel's inner address.

There are several reasons why this transient tunnel is established between the NAR and the mobile in the old PoA, unlike transient tunnel in FMIPv6 (Fast MIPv6) [RFC5568], where it is set up between mobile's new point of attachment and the old access router.

In case of inter-domain handoff, it is important that any signaling message between nPoA and the mobile needs to be secured. This transient secured tunnel provides the desired functionality including the securing the proactive binding update and transient data between the end-points before the handover has taken place. Unlike proactive mode of FMIPv6, transient handover packets are not sent to PAR, and thus a tunnel between mobile's new point of attachment and old access router is not needed.

In case of inter-domain handoff, PAR and NAR could logically be far from each other. Thus, the signaling and data during pre-authentication period will take a longer route, and thus, may be subjected to longer one-way-delay. Hence, MPA provides a tradeoff between larger packet loss or larger one-way-packet delay for a transient period, when the mobile is preparing to handoff.

The proactive handover tunnel is established using a tunnel management protocol. When IKEv2 is used for proactive IP address acquisition, IKEv2 is also used as the tunnel management protocol. Alternatively, when PANA is used for proactive IP address acquisition, PANA may be used as the secure tunnel management protocol.

Once the proactive handover tunnel is established between the mobile node and the access router in the candidate target network, the access router also needs to perform proxy address resolution (Proxy ARP) on behalf of the mobile node so that it can capture any packets

destined to the mobile node's new address.

Since the mobile needs to be able to communicate with the correspondent node while in the previous network some or all parts of binding update and data from the correspondent node to mobile node need to be sent back to the mobile node over a proactive handover tunnel. Details of these binding update procedure are described in Section 7.5.

In order for the traffic to be directed to the mobile node after the mobile node attaches to the target network, the proactive handover tunnel needs to be deleted or disabled. The tunnel management protocol used for establishing the tunnel is used for this purpose. Alternatively, when PANA is used as the authentication protocol the tunnel deletion or disabling at the access router can be triggered by means of PANA update mechanism as soon as the mobile moves to the target network. A link-layer trigger ensures that the mobile node is indeed connected to the target network and can also be used as the trigger to delete or disable the tunnel. A tunnel management protocol also triggers the router advertisement (RA) the from next access router to be sent over the tunnel, as soon as the tunnel creation is complete.

7.5. Binding Update

There are several kinds of binding update mechanisms for different mobility management schemes.

In case of Mobile IPv4 and Mobile IPv6, the mobile performs a binding update with the home agent only, if route optimization is not used. Otherwise, the mobile performs binding update with both the home agent (HA) and corresponding node (CN).

In case of SIP-based terminal mobility, the mobile sends binding update using INVITE to the correspondent node and REGISTER message to the Registrar. Based on the distance between the mobile and the correspondent node, the binding update may contribute to the handover delay. SIP-fast handover [SIPFAST] provides several ways of reducing the handover delay due to binding update. In case of secure proactive handover using SIP-based mobility management we do not encounter the delay due to binding update completely, as it takes place in the previous network.

Thus, this proactive binding update scheme looks more attractive when the correspondent node is too far from the communicating mobile node. Similarly, in case of Mobile IPv6, the mobile sends the newly acquired CoA from the target network as the binding update to the HA and CN. Also all signaling messages between MN and HA and between MN

and CN are passed through this proactive tunnel that is set up. These messages include Binding Update (BU), Binding Acknowledgement (BA) and the associated return routability messages such as Home Test Init (HoTI), Home Test (HoT), Care-of Test Init (CoTI), Care-of Test (COT). In Mobile IPv6, since the receipt of on-link router advertisement is mandatory for the mobile to detect the movement and trigger the binding update, router advertisement from next access router needs to be advertised over the tunnel. By proper configuration on NAR, router advertisement can be sent over the tunnel interface to trigger the proactive binding update. The mobile also needs to make the tunnel interface the active interface, so that it can send the binding update using this interface as soon as it receives the router advertisement.

If the proactive handover tunnel is realized as an IPsec tunnel, it will also protect these signaling messages between the tunnel end points and will make the return routability test secured as well. Any subsequent data will also be tunneled through as long as the mobile is in the previous network. The accompanying document [mpa-wireless] talks about the details of how binding updates and signaling for return routability are sent over the secured tunnel.

7.6. Preventing packet loss

7.6.1. Packet loss prevention in single interface MPA

For single interface MPA, there may be some transient packets during link-layer handover that are directed to the mobile node at the old point of attachment before the mobile node is able to attach to the target network. Those transient packets can be lost. Buffering these packets at the access router of the old point of attachment can eliminate packet loss. Dynamic buffering signals that are signalled from the MN can temporarily hold transient traffic during handover and then these packets can be forwarded to the MN once it attaches to the target network. A detailed analysis of buffering technique can be found in [PIMRC06].

An alternative method is to use bicasting. Bicasting helps to forward the traffic to two destinations at the same time. However, it does not eliminate packet loss if link-layer handover is not seamlessly performed. On the other hand, buffering does not reduce packet delay. While packet delay can be compensated by a playout buffer at the receiver side for streaming application, a playout buffer does not help much for interactive VoIP application that cannot tolerate for large delay jitters. Thus it is still important to optimize the link-layer handover anyway.

7.6.2. Preventing packet losses for multiple interfaces

MPA usage in multi-interface handover scenarios involves preparing the second interface for use via the current active interface. This preparation involves pre-authentication and provisioning at a target network where the second interface would be the eventual active interface. For example, during inter-technology handover from a Wi-Fi to a CDMA network, pre-authentication at the CDMA network can be performed via the Wi-Fi interface. The actual handover occurs when the CDMA interface becomes the active interface for the MN.

In such scenarios, if handover occurs while both interfaces are active, there is generally no packet loss since transient packets directed towards the old interface will still reach the MN. However, if sudden disconnection of the current active interface is used to initiate handover to the prepared interface then transient packets for the disconnected interface will be lost while the MN attempts to be reachable at the prepared interface. In such cases, a specialized form of buffering can be used to eliminate packet loss where packets are merely copied at an access router in the current active network prior to disconnection. If sudden disconnection does occur, copied packets can be forwarded to the MN once the prepared interface becomes the active reachable interface. The copy-and-foward mechanism is not limited to multi-interface handover.

A notable side-effect of this process is the possible duplication of packets during forwarding to the new active interface. Several approaches can be employed to minimize this effect. Relying on upper layer protocols such as TCP to detect and eliminate duplicates is the most common approach. Customized duplicate detection and handling techniques can also be used. In general, packet duplication is a well known issue that can also be handled locally by the MN.

If the mobile takes a longer amount of time to detect the disconnection event of the current active interface, it can also have an adverse effect on the length of the handover process. Thus it becomes necessary to use an optimized scheme of detecting interface disconnection in such scenarios. Use of the current interface to perform pre-authentication instead of the new interface is desirable in certain circumstances, such as to save battery power or in cases where the adjacent cells (e.g., WiFi, and CDMA) are non-overlapping or in cases when the carrier does not allow simultaneous use of both interfaces. However, in certain circumstances, depending upon the type of target network, only parts of MPA operations can be performed (e.g., pre-authentication, pre-configuration, proactive binding update). In a specific scenario involving handoff between WiFi and CDMA network, some of the PPP context can be set up during the pre-authentication period, thus reducing the time for PPP activation.

7.6.3. Reachability test

In addition to previous techniques, the MN may also want to ensure reachability of the new point of attachment before switching from the old one. This can be done by exchanging link-layer management frames with the new point of attachment. This reachability check should be performed as quickly as possible. In order to prevent packet loss during this reachability check, transmission of packets over the link between the MN and old point of attachment should be suspended by buffering the packets at both ends of the link during the reachability check. How to perform this buffering is out of scope of this document. Some of the results using this buffering scheme are explained in the accompanying implementation document.

7.7. Security and mobility

7.7.1. Link-layer security and mobility

Using the MPA-SA established between the mobile node and the authentication agent for a CTN, during the pre-authentication phase, it is possible to bootstrap link-layer security in the CTN while the mobile node is in the current network in the following way. Figure 5 shows the sequence of operation.

(1) The authentication agent and the mobile node derives a PMK (Pair-wise Master Key) [RFC5247] using the MPA-SA that is established as a result of successful pre-authentication. Successful operation of EAP and an AAA protocol may be involved during pre-authentication to establish the MPA-SA. From the PMK, distinct TSKs (Transient Session Keys) [RFC5247] for the mobile node are directly or indirectly derived for each point of attachment of the CTN.

(2) The authentication agent may install the keys derived from the PMK and used for secure association to points of attachment. The derived keys may be TSKs or intermediary keys from which TSKs are derived.

(3) After the mobile node chooses a CTN as the target network and switches to a point of attachment in the target network (which now becomes the new network for the mobile node), it executes a secure association protocol such as the IEEE 802.11i 4-way handshake [802.11] using the PMK in order to establish PTKs (Pair-wise Transient Keys) and GTKs (Group Transient Keys) [RFC5247] used for protecting link-layer packets between the mobile node and the point of attachment. No additional execution of EAP authentication is needed here.

(4) While the mobile node is roaming in the new network, the mobile

node only needs to perform a secure association protocol with its point of attachment point and no additional execution of EAP authentication is needed either. Integration of MPA with link-layer handover optimization mechanisms such as 802.11r can be archived this way.

The mobile node may need to know the link-layer identities of the point of attachments in the CTN to derive TSKs.

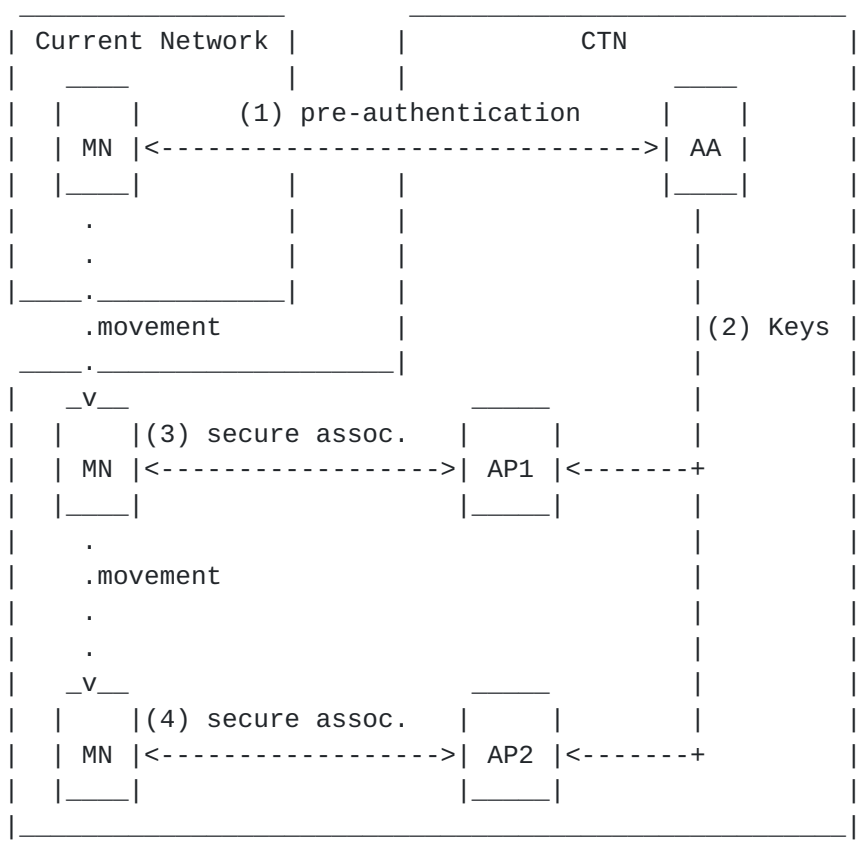

Figure 5: Bootstrapping Link-layer Security

7.7.2. IP layer security and mobility

IP layer security is typically maintained between the mobile and first hop router or any other network element such as SIP proxy by means of IPsec. This IPSec SA can be set up either in tunnel mode or in ESP mode. However, as the mobile moves, the IP address of the router and outbound proxy will change in the new network. The mobile's IP address may or may not change depending upon the mobility protocol being used. This will warrant re-establishing a new security association between the mobile and the desired network entity. In some cases such as in 3GPP/3GPP2 IMS/MMD environment data traffic is not allowed to pass through unless there is an IPsec SA

established between the mobile and outbound proxy. This will of course add unreasonable delay to the existing real-time communication during mobile's movement. In this scenario, key exchange is done as part of SIP registration that follows a key exchange procedure called AKA (Authentication and Key Agreement).

MPA can be used to bootstrap this security association as part of pre-authentication via the new outbound proxy. Prior to the movement, if the mobile can pre-register via the new outbound proxy in the target network and completes the pre-authentication procedure, then the new SA state between the mobile and new outbound proxy can be established prior to the movement to the new network. A similar approach can also be applied if a key exchange mechanism other than AKA is used or the network element with which the security association has to be established is different than an outbound proxy.

By having the security association established ahead of time, the mobile does not need to involve in any exchange to set up the new security association after the movement. Any further key exchange will be limited to renew the expiry time. This will also reduce the delay for real-time communication as well.

7.8. Authentication in initial network attachment

When the mobile node initially attaches to a network, network access authentication would occur regardless of the use of MPA. The protocol used for network access authentication when MPA is used for handover optimization can be a link-layer network access authentication protocol such as IEEE 802.1X or a higher-layer network access authentication protocol such as PANA.

8. Security Considerations

This document describes a framework of a secure handover optimization mechanism based on performing handover-related signaling between a mobile node and one or more candidate target networks to which the mobile node may move in the future. This framework involves acquisition of the resources from the CTN as well as data packet redirection from the CTN to the mobile node in the current network before the mobile node physically connects to one of those CTN.

Acquisition of the resources from the candidate target networks must be done with appropriate authentication and authorization procedures in order to prevent an unauthorized mobile node from obtaining the resources. For this reason, it is important for the MPA framework to perform pre-authentication between the mobile node and the candidate

target networks. The MN-CA key and the MN-AR key generated as a result of successful pre-authentication can protect subsequent handover signaling packets and data packets exchanged between the mobile node and the MPA functional elements in the CTN's.

The MPA framework also addresses security issues when the handover is performed across multiple administrative domains. With MPA, it is possible for handover signaling to be performed based on direct communication between the mobile node and routers or mobility agents in the candidate target networks. This eliminates the need for a context transfer protocol [RFC5247] for which known limitations exist in terms of security and authorization. For this reason, the MPA framework does not require trust relationship among administrative domains or access routers, which makes the framework more deployable in the Internet without compromising the security in mobile environments.

9. IANA Considerations

This document has no actions for IANA.

10. Acknowledgments

We would like to thank Farooq Anjum and Raziq Yaqub for their review of this document, and Subir Das for standardization support in the IEEE 802.21 WG.

Authors would like to acknowledge Christian Vogt, Rajeev Koodli, Marco Liebsch, Juergen Schoenwaelder and Charles Perkins for their thorough review of the draft and useful feedback.

11. References

11.1. Normative References

[RFC5944] Perkins, C., "IP Mobility Support for IPv4, Revised", RFC 5944, November 2010.

[RFC3748] Aboba, B., Blunk, L., Vollbrecht, J., Carlson, J., and H. Levkowetz, "Extensible Authentication Protocol (EAP)", RFC 3748, June 2004.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

[RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC5380] Soliman, H., Castelluccia, C., ElMalki, K., and L. Bellier, "Hierarchical Mobile IPv6 (HMIPv6) Mobility Management", RFC 5380, October 2008.

[RFC5568] Koodli, R., "Mobile IPv6 Fast Handovers", RFC 5568, July 2009.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4555] Eronen, P., "IKEv2 Mobility and Multihoming Protocol (MOBIKE)", RFC 4555, June 2006.

[RFC4881] El Malki, K., "Low-Latency Handoffs in Mobile IPv4", RFC 4881, June 2007.

[RFC4066] Liebsch, M., Singh, A., Chaskar, H., Funato, D., and E. Shim, "Candidate Access Router Discovery (CARD)", RFC 4066, July 2005.

[RFC4830] Kempf, J., "Problem Statement for Network-Based Localized Mobility Management (NETLMM)", RFC 4830, April 2007.

[RFC4831] Kempf, J., "Goals for Network-Based Localized Mobility Management (NETLMM)", RFC 4831, April 2007.

[RFC4065] Kempf, J., "Instructions for Seamoby and Experimental Mobility Protocol IANA Allocations", RFC 4065, July 2005.

[RFC5247] Aboba, B., Simon, D., and P. Eronen, "Extensible Authentication Protocol (EAP) Key Management Framework", RFC 5247, August 2008.

[RFC5191] Forsberg, D., Ohba, Y., Patil, B., Tschofenig, H., and A. Yegin, "Protocol for Carrying Authentication for Network Access (PANA)", RFC 5191, May 2008.

[RG98] ITU-T, "General Characteristics of International Telephone Connections and International Telephone Circuits: One-Way Transmission Time", ITU-T Recommendation G.114 1998.

[ITU98] ITU-T, "The E-Model, a computational model for use in transmission planning", ITU-T Recommendation G.107 1998.

[ETSI] ETSI, "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3: End-to-end Quality of Service in TIPHON systems; Part 1: General aspects of Quality of Service.", ETSI TR 101 329-6 V2.1.1.

11.2. Informative References

[RFC5201] Moskowitz, R., Nikander, P., Jokela, P., and T. Henderson, "Host Identity Protocol", RFC 5201, April 2008.

[RFC2679] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.

[RFC2680] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Packet Loss Metric for IPPM", RFC 2680, September 1999.

[RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, September 1999.

[RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

[RFC2608] Guttman, E., Perkins, C., Veizades, J., and M. Day, "Service Location Protocol, Version 2", RFC 2608, June 1999.

[RFC2473] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[RFC3046] Patrick, M., "DHCP Relay Agent Information Option", RFC 3046, January 2001.

[RFC4039] Park, S., Kim, P., and B. Volz, "Rapid Commit Option for the Dynamic Host Configuration Protocol version 4 (DHCPv4)", RFC 4039, March 2005.

[RFC5172] Varada, S., "Negotiation for IPv6 Datagram Compression Using IPv6 Control Protocol", RFC 5172, March 2008.

[RFC5648] Wakikawa, R., Devarapalli, V., Tsirtsis, G., Ernst, T., and K. Nagami, "Multiple Care-of Addresses Registration", RFC 5648, October 2009.

[RFC4429] Moore, N., "Optimistic Duplicate Address Detection (DAD) for IPv6", RFC 4429, April 2006.

[RFC5836] Ohba, Y., Wu, Q., and G. Zorn, "Extensible Authentication Protocol (EAP) Early Authentication Problem Statement",

RFC 5836, April 2010.

[RFC5213] Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, August 2008.

[RFC5974] Manner, J., Karagiannis, G., and A. McDonald, "NSIS Signaling Layer Protocol (NSLP) for Quality-of-Service Signaling", RFC 5974, October 2010.

[RFC5169] Clancy, T., Nakhjiri, M., Narayanan, V., and L. Dondeti, "Handover Key Management and Re-Authentication Problem Statement", RFC 5169, March 2008.

[SIPMM] Schulzrinne, H. and E. Wedlund, "Application Layer Mobility Using SIP", ACM MC2R.

[CELLIP] Cambell, A., Gomez, J., Kim, S., Valko, A., and C. Wan, "Design, Implementation, and Evaluation of Cellular IP", IEEE Personal communication Auguest 2000.

[MOBIQUIT07] Lopez, R., Dutta, A., Ohba, Y., Schulzrinne, H., and A. Skarmeta, "Network-layer assisted mechanism to optimize authentication delay during handoff in 802.11 networks", IEEE Mobiquitous June 2007.

[IEEE-03-084] Mishra, A., Shin, M., Arbaugh, W., Lee, I., and K. Jang, "Proactive Key Distribution to support fast and secure roaming, IEEE 802.11 Working Group, IEEE-03-084r1-I, "www.ieee802.org/11/Documents/DocumentHolder/3-084.zip"", IEEE June 2003.

[SPRINGER07] Dutta, A., Das, S., Famolari, D., Ohba, Y., Taniuchi, K., Fajardo, V., Schulzrinne, H., Lopez, R., Kodama, T., and A. Skarmeta, "Seamless proactive handover across heterogeneous access networks", Wireless Personal Communication February 2007.

[HAWAII] Ramjee, R., Porta, T., Thuel, S., Varadhan, K., and S. Wang, "HAWAII: A Domain-based Approach for Supporting Mobility in Wide-area Wireless networks", International Conference on Network Protocols ICNP'99.

[IDMP] Das, S., Dutta, A., Misra, A., and S. Das, "IDMP: An Intra-Domain Mobility Management Protocol for Next Generation Wireless Networks", IEEE Wireless Communication

Magazine October 2000.

[I-D.ietf-mobileip-reg-tunnel]
Calhoun, P., Montenegro, G., Perkins, C., and E. Gustafsson, "Mobile IPv4 Regional Registration", draft-ietf-mobileip-reg-tunnel-09 (work in progress), July 2004.

[YOKOTA] Yokota, H., Idoue, A., and T. Hasegawa, "Link Layer Assisted Mobile IP Fast Handoff Method over Wireless LAN Networks", Proceedings of ACM Mobicom 2002.

[MACD] Shin, S., "Reducing MAC Layer Handoff Latency in IEEE 802.11 Wireless LANs", MOBIWAC Workshop .

[SUM] Dutta, A., Zhang, T., Madhani, S., Taniuchi, K., Ohba, Y., and H. Schulzrinne, "Secured Universal Mobility", WMASH 2004.

[SIPFAST] Dutta, A., Madhani, S., and H. Schulzrinne, "Fast handoff Schemes for Application Layer Mobility Management", PIMRC 2004.

[PIMRC06] Dutta, A., Ohba, Y., and H. Schulzrinne, "Dynamic Buffering Protocol for Mobile", PIMRC 2006.

[MITH] Gwon, Y., Fu, G., and R. Jain, "Fast Handoffs in Wireless LAN Networks using Mobile initiated Tunneling Handoff Protocol for IPv4 (MITHv4)", Wireless Communications and Networking 2003, January 2005.

[Wenyu] Jiang, W. and H. Schulzrinne, "Modeling of Packet Loss and Delay and their Effect on Real-Time Multimedia Service Quality", NOSSDAV 2000, June 2000.

[Romdhani]
Romdhani, I., Kellil, M., Lach, H., and A. Bouabdallah, "IP Mobile Multicast Challenges and Solutions", IEEE Communication Magazine 2004, March 2000.

[802.21] "IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services, IEEE 802.21-2008", A contribution to IEEE 802.21 WG , January 2009.

[802.11] "IEEE Wireless LAN Edition A compilation based on IEEE Std 802.11-1999(R2003)", Institute of Electrical and Electronics Engineers September 2003.

[GPSIP] Dutta, A., "GPS-IP based fast-handoff for Mobiles", IEEE Sarnoff Symposium 2006.

[MAGUIRE] Vatn, J. and G. Maguire, "The effect of using co-located care-of-address on macro handover latency", 14th Nordic Teletraffic Seminar 1998.

[mpa-mobike] Mghazli, Y. and J. Bournelle, "MPA using IKEv2 and MOBIKE", draft-yacine-preauth-ipsec-00 IETF.

[FMIP-results] Cabellos-Apaicio, A., Nunez-Martinez, J., Julian-Bertomeu, H., Jakab, L., Serral-Gracia, R., and J. Domingo-Pascual, "Evaluation of Fast Handover Implementation for Mobile IPv6 in a Real Testbed", IPOM 2005 LNCS 3751.

[mpa-wireless] Dutta, A., Famolari, D., Das, S., Ohba, Y., Fajardo, V., Taniuchi, K., Lopez, R., and H. Schulzrinne, "Media-Independent Pre-authentication Supporting Secure Interdomain Handover Optimization", IEEE Wireless Magazine April 2008.

Appendix A. Proactive duplicate address detection

When the DHCP server dispenses an IP address, it updates its lease table, so that this same address is not given to another client for that specific period of time. At the same time the client also keeps a lease table locally so that it can renew when needed. In some cases where a network consists of both DHCP and non-DHCP enabled clients, there is a probability that another client in the LAN may have been configured with an IP address from the DHCP address pool. In such scenario the server detects a duplicate address based on ARP (Address Resolution Protocol) or IPv6 Neighbor Discovery before assigning the IP address. This detection procedure may take from 4 sec to 15 sec [MAGUIRE] and will thus contribute to a larger handover delay. In case of a proactive IP address acquisition process, this detection is performed ahead of time and thus, does not affect the handover delay at all. By performing the duplicate address detection ahead of time, we reduce the IP address acquisition time.

The proactive duplicate address detection (DAD) over the candidate target network should be performed by the NAR on behalf of the mobile at the time of proactive handover tunnel establishment since duplicate address detection over a tunnel is not always performed. For example, in the case of IPv6, DAD over an IP-IP tunnel interface

is turned off in an existing implementation. In the case of IPv6 over PPP [RFC5172], IPCPv6 negotiates the link local addresses and hence DAD over the tunnel is not needed. After the mobile has moved to the target network, a DAD procedure may be started because of reassignment of the nCoA to the physical interface to the target network. In that case, the mobile should use optimistic DAD [RFC4429] over the physical interface so that the nCoA that was used inside the proactive handover tunnel before handover can be immediately used over that physical interface after handover. The schemes used for the proactive DAD and optimistic DAD are applicable to both stateless and stateful address autoconfiguration schemes used for obtaining a nCoA.

Appendix B. Address resolution

Address resolution involves updating next access router's neighbor cache. We briefly describe these two operations below.

During the process of pre-configuration, the MAC address resolution mappings needed by the mobile node to communicate with nodes in the target network after attaching to the target network can also be known, where the communicating nodes maybe the access router, authentication agent, configuration agent and correspondent node. There are several possible ways of performing such proactive MAC address resolution.

- Use an information service mechanism [802.21] to resolve the MAC addresses of the nodes. This might require each node in the target network to be involved in the information service so that the server of the information service can construct the database for proactive MAC address resolution.

- Extend the authentication protocol used for pre-authentication or the configuration protocol used for pre-configuration to support proactive MAC address resolution. For example, if PANA is used as the authentication protocol for pre-authentication, PANA messages may carry AVPs used for proactive address resolution. In this case, the PANA authentication agent in the target network may perform address resolution for on behalf of the mobile node.

- One can also make use of DNS to map the MAC address of the specific interface associated with a specific IP address of the network element in the target network. One may define a new DNS resource record (RR) to proactively resolve the MAC addresses of the nodes in the target network. But this approach may have its own limitations since a MAC address is a resource that is bound to an IP address, not directly to a domain name.

When the mobile node attaches to the target network, it installs the proactively obtained address resolution mappings without necessarily performing address resolution queries for the nodes in the target network.

On the other hand, the nodes that reside in the target network and are communicating with the mobile node should also update their address resolution mappings for the mobile node as soon as the mobile node attaches to the target network. The above proactive address resolution methods could also be used for those nodes to proactively resolve the MAC address of the mobile node before the mobile node attaches to the target network. However, this is not useful since those nodes need to detect the attachment of the mobile node to the target network before adopting the proactively resolved address resolution mapping. A better approach would be integration of attachment detection and address resolution mapping update. This is based on gratuitously performing address resolution [RFC5944], [RFC3775] in which the mobile node sends an ARP Request or an ARP Reply in the case of IPv4 or a Neighbor Advertisement in the case of IPv6 immediately after the mobile node attaches to the new network so that the nodes in the target network can quickly update the address resolution mapping for the mobile node.

Appendix C. MPA Deployment Issues

In this section we describe some of the deployment issues related to MPA.

C.1. Considerations for failed switching and switch-back

The ping-Pong effect is one of the common problems found during handover. The Ping-pong effect arises when a mobile is located at the borderline of the cell or decision point and a handover procedure is frequently executed. This results in higher call drop probability, lower connection quality, increased signaling traffic and waste of resources. All of these affect mobility optimization. Handoff algorithms are the deciding factors for performing the handoff between the networks. Traditionally these algorithms employ a threshold to compare the values of different metrics to decide on the handoff. These metrics include signal strength, path loss, carrier-to-interference ratios (CIR), Signal to Interference Ratios (SIR), Bit Error Rate (BER), power budget. In order to avoid the ping-pong effect, some additional parameters are employed by the decision algorithm such as hystereris margin, dwell timers, and averaging window. For a vehicle moving with a high speed, other parameters such as distance between the mobile node and the point of attachment, velocity of the mobile, location of the mobile, traffic

and bandwidth characteristics are also taken into account to reduce the ping-pong effect. Most recently there are other handoff algorithms that help reduce the ping-pong effect in a heterogeneous network environment that are based on techniques such as hypothesis testing, dynamic programming and pattern recognition techniques. While it is important to devise smart handoff algorithms to reduce the ping-pong effect, it is also important to devise methods to recover from this effect.

In the case of the MPA framework, the ping-pong effect will result in the back-and-forth movement of the mobile between the current network and target network and between the candidate target networks. MPA in its current form will be affected because of a number of tunnels setup between the mobile and neighboring access routers, number of binding updates and associated handoff latency resulting out of ping-pong situation. The mobile's handoff rate may also contribute to delay and packet loss. We propose few techniques that will help reduce the probability of ping-pong and propose several methods for the MPA framework so that it can recover from the packet loss resulting out of the ping-pong effect.

The MPA framework can take advantage of the mobile's geo-location with respect to APs in the neighboring networks using GPS. In order to avoid the oscillation between the networks, a location-aware algorithm can be derived by using a co-relation between user's location and cached data from the previous handover attempts. In some cases only location may not be the only indicator for a handoff decision. For example in Manhattan type grid networks, although a mobile is close to an AP, it may not have enough SNR (Signal to Noise Ration) to make a good connection. Thus knowledge of mobility pattern, dwell time in a call and path identification will help avoid the ping-pong problem to a great extent.

In the absence of a good handoff algorithm that can avoid ping-pong effect, it may be required to put in place a good recovery mechanism so as to mitigate the effect of ping-pong. It may be necessary to keep the established context in the current network for a period of time, so that it can be quickly recovered when the mobile comes back to the network where the context was last used. This context may include security association, IP address used, tunnels established. Bicasting the data to both the previous network and the new network for a predefined period will also help the mobile to take care of the lost packets in case the mobile moves back and forth between the networks. The mobile can also take certain action, after it determines that it is in a stable state with respect to a ping-pong situation.

When the MPA framework takes advantage of a combination of IKEv2 and

MOBIKE, the ping-pong effect can be reduced further [mpa-mobike].

C.2. Authentication state management

In case of pre-authentication with multiple target networks, it is useful to maintain the state in the authentication agent of each of the neighboring networks for certain time. Thus, if the mobile does move back and forth between neighboring networks, already maintained authentication state can be helpful. We provide some highlights on multiple security association state management below.

A mobile node that has pre-authenticated with an authentication agent in a candidate target network and has a MPA-SA, may need to continue to keep the MPA-SA while it continues to stay in the current network or even after it does handover to a network that is different from the candidate target network.

When an MN that has been authenticated and authorized by an authentication agent in the current network makes a handover to a target network, it may want to hold the SA that has been established between the MN and the authentication agent for a certain time period so that it does not have to go through the entire authentication signaling to create an SA from scratch in case it returns to the previous network. Such an SA being held at the authentication agent after the MN's handover to other network is considered as an MPA-SA. In this case, the authentication agent should change the fully authorized state for the MN to an unauthorized state. The unauthorized state can be changed to the fully authorized state only when the MN comes back to the network and provides a proof of possession of a key associated with the MPA-SA.

While an MPA-SA is being held at an authentication agent, the MN will need to keep updating the authentication agent when an IP address of the MN changes due to a handover to re-establish the new SA.

C.3. Pre-allocation of QoS resources

In the pre-configuration phase, it is also possible to pre-allocate QoS resources that may be used by the mobile node not only after handover but also before handover. When pre-allocated QoS resources are used before handover, it is used for application traffic carried over a proactive handover tunnel.

It is possible that QoS resources are pre-allocated in an end-to-end fashion. One method to achieve this proactive end-to-end QoS reservation is to execute NSLP [RFC5974] or RSVP [RFC2205] over a proactive handover tunnel where pre-authentication can be used for bootstrapping a security association for the proactive handover

tunnel to protect the QoS signaling. In this case, QoS resources are pre-allocated on the path between the correspondent node and a target access router can be used continuously before and after handover. On the other hand, duplicate pre-allocation of QoS resources between the target access router and the mobile node is necessary when using pre-allocated QoS resources before handover due to difference in paths between the target access router and the mobile node before and after handover. QoS resources to be used for the path between the target access router and the mobile node after handover may be pre-allocated by extending NSLP to work for off-path signaling (Note: this path can be viewed as off-path before handover) or by media-specific QoS signaling at layer 2.

C.4. Resource allocation issue during pre-authentication

In case of multiple CTNs, establishing multiple tunnels with the neighboring target networks provides some additional benefits. But it also contributes to some resource utilization issues as well. A pre-authentication process with multiple candidate target networks can happen in several ways.

The very basic scheme involves authenticating the mobile with the multiple authentication agents in the neighboring networks, but actual pre-configuration and binding update take place only after layer 2 movement to a specific network is complete.

Similarly, in addition to pre-authentication, the mobile can also complete the pre-configuration while in the previous network, but can postpone the binding update until after the mobile has moved. Like the previous case, in this case the mobile also does not need to set up the pre-configured tunnels. While the pre-authentication process and part of the pre-configuration process are taken care of before the mobile has moved to the new network, binding update is actually done after the mobile has moved.

The third type of multiple pre-authentication involves all the three steps while the mobile is in the previous networks, such as authentication, configuration and binding update. But, this specific process utilizes the most amount of resources. Some of the resources that get used during this process are as follows:

1)Additional signaling for pre-authentication in the neighboring networks

2)Holding the IP address of the neighboring networks in mobiles cache for certain amount of time. It needs additional processing in the mobile for storing these IP addresses. In addition it also uses up the temporary IP addresses from the neighboring routers.

3)There is an additional cost associated with setting up additional transient tunnels with the target routers in the neighboring networks and mobile.

4) In case of binding update with multiple IP addresses obtained from the neighboring networks, multiple transient streams flow between the CN and mobile using these transient tunnels.

When only pre-authentication and pre-configuration are done ahead of time with multiple networks, the mobile sends one binding update to the CN. In this case it is important to find out when to send the binding update after the layer 2 handoff.

In case binding update with multiple contact addresses is sent, multiple media streams stem out of CN using the transient tunnels. But in that case one needs to send another Binding Update after the handover with the contact address set to the new address (only one address) where the mobile has moved. This way the mobile stops sending media to other neighboring networks where the mobile did not move.

The following is an illustration of this specific case that takes care of multiple binding streams, when the mobile moves only to a specific network, but sends multiple binding updates in the previous network. MN sends a binding update to CH with multiple contact addresses such as c1,c2, and c3 that were obtained from three neighboring networks. This allows the CN to send transient multiple streams to the mobile over the pre-established tunnels. After the mobile moves to the actual network, it sends another binding update to the CN with the care-of-address of the mobile in the network where the mobile has moved in. Some of the issues with multiple streams are consumption of extra bandwidth for a small period of time.

Alternatively, one can apply the buffering technique at the target access router or at the home agent. Transient data can be forwarded to the mobile after it has moved in. Forwarding of data can be triggered by the mobile either as part of Mobile IP registration or as a separate buffering protocol.

C.5. Systems evaluation and performance results

In this Section, we present some of the results from MPA implementation when applied to different handover scenarios. We present the summary of results from our experiments using MPA techniques for two types of handovers I) Intra-technology and Intra-domain, II) Inter-technology and Inter-domain. We also present the results from how MPA can bootstrap layer 2 security for both roaming and non-roaming cases. Detailed procedure and results are explained

in [MOBIQUIT07] and [SPRINGER07].

C.5.1. Intra-technology, Intra-domain

The results for MIPv6 and SIP mobility involving intra-domain mobility are shown in Figure 6 and Figure 7, respectively.

	Buffering (disabled) & RO (disabled)	Buffering (enabled) & RO (disabled)	Buffering (disabled) & RO (enabled)	Buffering (enabled) & RO (enabled)
L2 handoff (ms)	4.00	4.33	4.00	4.00
L3 handoff (ms)	1.00	1.00	1.00	1.00
Avg. packet loss	1.33	0	0.66	0
Avg. inter-packet arrival interval (ms)	16.00	16.00	16.00	16.00
Avg. inter-packet arrival time during handover (ms)	n/a	45.33	n/a	66.60
Avg. packet jitter (ms)	n/a	29.33	n/a	50.60
Buffering Period (ms)	n/a	50.00	n/a	50.00
Buffered Packets	n/a	2.00	n/a	3.00

Figure 6: Mobile IPv6 with MPA Results

	Buffering disabled	Buffering enabled
L2 handoff (ms)	4.00	5.00
L3 handoff (ms)	1.00	1.00
Avg. packet loss	1.50	0
Avg. inter-packet arrival interval (ms)	16.00	16.00
Avg. inter-packet arrival time during handover (ms)	n/a	29.00
Avg. packet jitter (ms)	n/a	13.00
Buffering Period (ms)	n/a	20.00
Buffered Packets	n/a	3.00

Figure 7: SIP Mobility with MPA Results

For all measurement, we did not experience any performance degradation during handover in terms of the audio quality of the voice traffic.

With the use of buffering during handover, packet loss during the actual L2 and L3 handover is eliminated with an appropriate and reasonable settings of the buffering period for both MIP6 and SIP mobility. In the case of MIP6, there is not a significant difference in results with and without route optimization. It should be noted that results with more samples would be necessary for a more detailed analysis.

In case of non-MPA assisted handover, handover delay and associated packet loss occurs from the moment the link-layer handover procedure begins up to successful processing of the binding update. During this process, IP address acquisitions via DHCP incurs the longest delay. This is due to the detection of duplicate IP address in the network before DHCP request completes. Binding update exchange also experiences long delay if the CN is too far from the MN. As a result, the Non-MPA assisted handover took an average of 4 seconds to

complete with an approximate packet loss of about 200 packets. The measurement is based on the same traffic rate and traffic source as the MPA assisted handover.

C.5.2. Inter-technology, Inter-domain

Handoff involving heterogeneous access can take place in many different ways. We limit the experiment to two interfaces and therefore results in several possible setup scenarios depending upon the activity of the second interface. In one scenario, the second interface comes up when the link to the first interface goes down. This is a reactive scenario and usually gives rise to undesirable packet loss and handoff delay. In a second scenario, the second interface is being prepared while the mobile still communicates using the old interface. Preparation of the second interface should include setup of all the required state and security associations (e.g., PPP state, LCP, CHAP). If such lengthly process is established ahead of time, it reduces the time taken for the secondary interface to be attached to the network. After preparation, the mobile decides to use the second interface as the active interface. This results in less packet loss as it uses make-before-break techniques. This is a proactive scenario and can have two flavors. The first is where both interfaces are up and the second is when only the old interface is up the prepared interface is brought up only when handoff is about to occur. This scenario may be beneficial from a battery management standpoint. Devices that operate two interfaces simultaneously can rapidly deplete their batteries. However, by activating the second interface only after an appropriate network has been selected the client may utilize battery effectively.

As compared to non-optimized handover that may result in delay up to 18 sec and 1000 packet loss during handover from WLAN to CDMA, we observed 0 packet loss, and 50 ms handoff delay between the last pre-handoff packet and first in-handoff packet. This handoff delay includes the time due to link down detection and time needed to delete the tunnel after the mobile has moved. However, we observed about 10 duplicate packets because of the copy-and-forward mechanism at the access routers. But these duplicate packets are usually handled easily by the upper layer protocols.

C.5.3. MPA-assisted Layer 2 pre-authentication

In this section, we discuss the results obtained from MPA-assisted layer 2 pre-authentication and compare these with EAP authentication and IEEE 802.11i's pre-authentication techniques. Figure 12 shows the experimental testbed where we have conducted the MPA-assisted pre-authentication experiment for bootstrapping layer 2 security as

explained in Section 7. By pre-authenticating and pre-configuring the link, the security association procedure during handoff reduces to a 4-way handshake only. Then MN moves to the AP and, after association, runs a 4-way handshake by using the PSKap (Pre-shared Key at AP) generated during PANA pre-authentication. At this point the handoff is complete. Details of this experimental testbed can be found in [MOBIQUIT07].

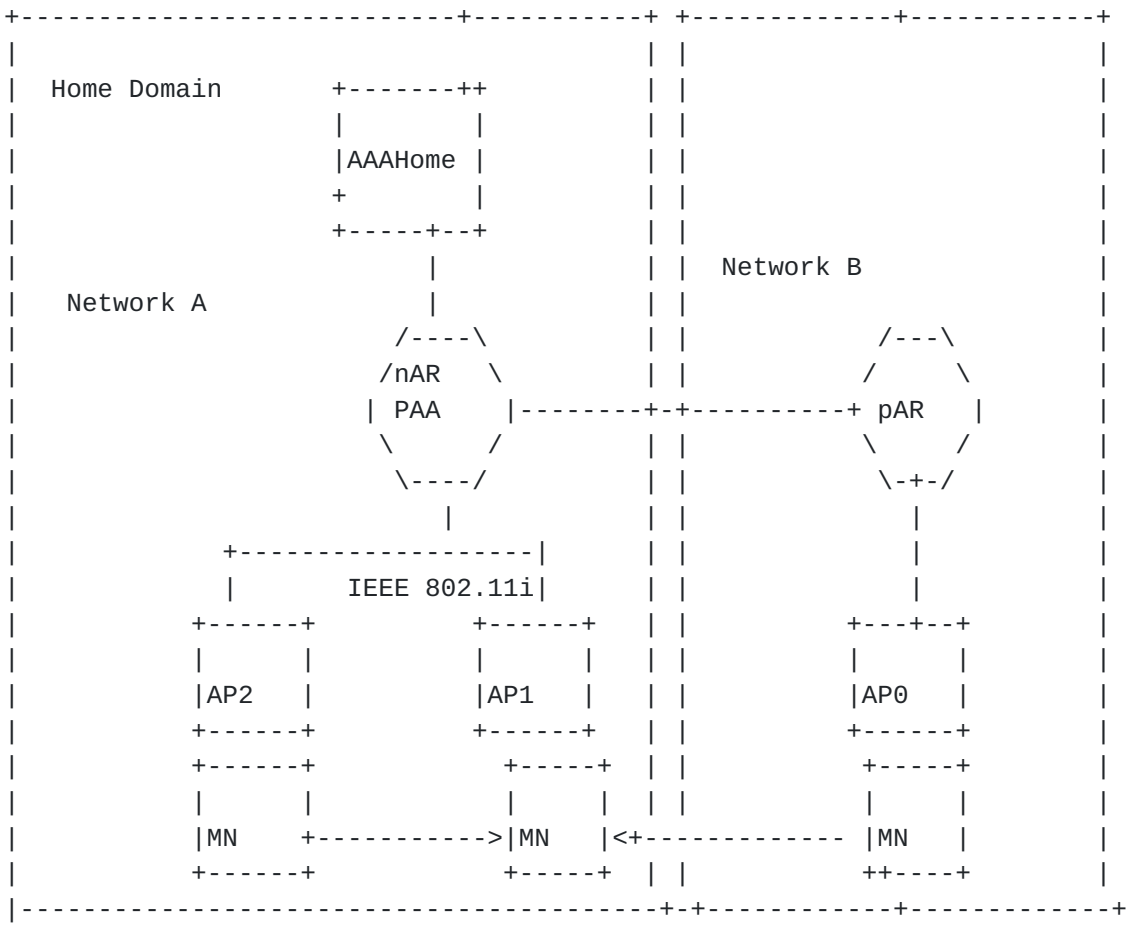

Figure 8: Experimental Testbed for MPA-assisted L2 Pre-authentication (Non-roaming)

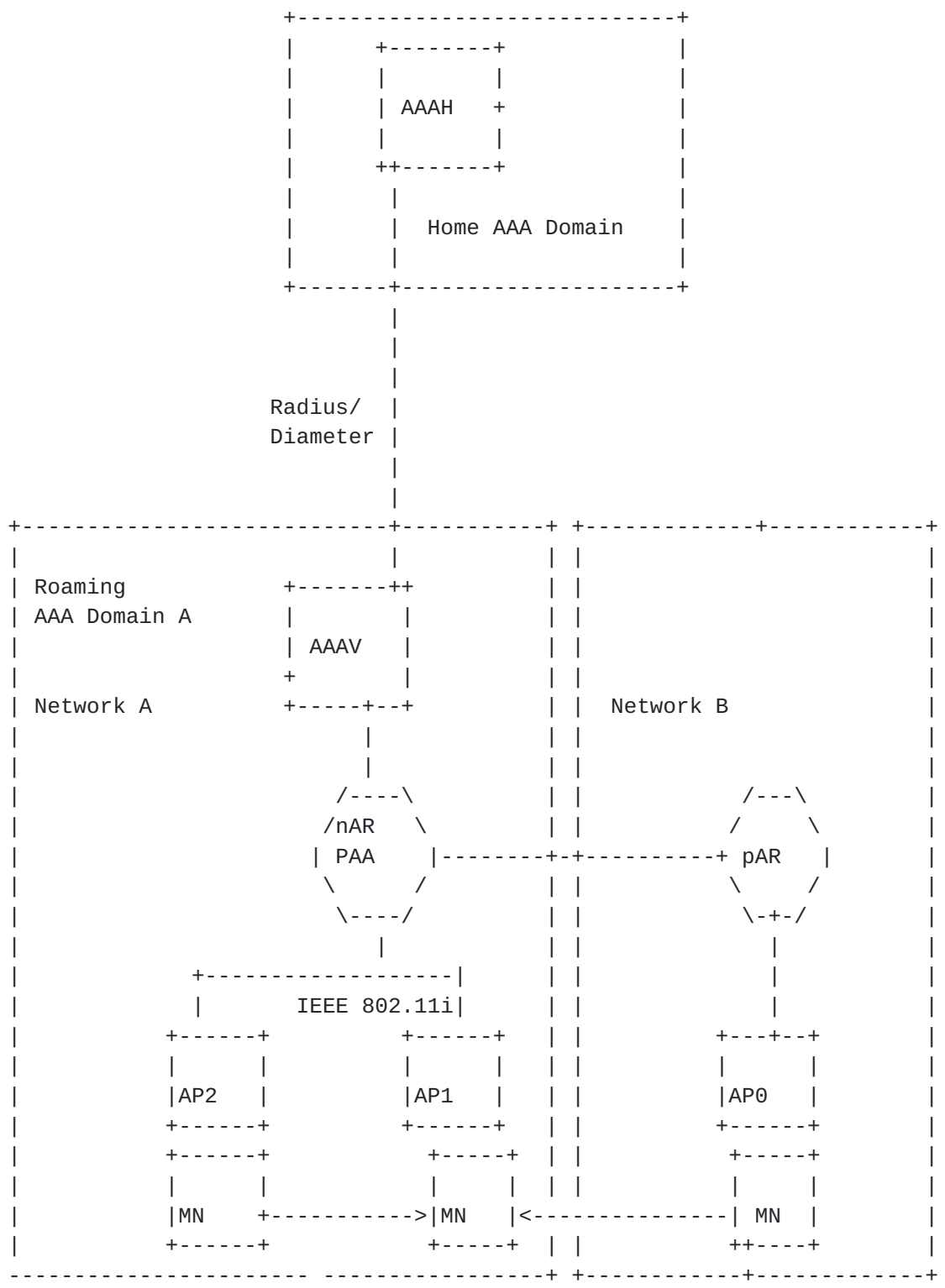

Figure 9: Experimental Testbed for MPA-assisted L2 Pre-authentication (Roaming)

We have experimented with three types of movement scenarios involving both non-roaming and roaming cases using the testbeds shown in figures 12 and 13, respectively. In the roaming case, MN is visiting in a domain different than its home domain. Consequently, the AAAh needs to be contacted which is placed in a location far from the visiting domain. For the non-roaming case, we assume the MN is moving within its home domain and only the local AAA server (AAAHome) is contacted which is the home AAA server for the mobile.

The first scenario does not involve any pre-authentication. The MN is initially connected to AP0 and moves to AP1. Because neither network-layer authentication is enabled nor IEEE 802.11i pre-authentication is used, the MN needs to engage in a full EAP authentication with AP1 to gain access to the network after the move (post-authentication). This experiment shows the effect of absence of any kind of pre-authentication.

The second scenario involves 802.11i pre-authentication and involves movement between AP1 and AP2. In this scenario, the MN is initially connected to AP2, and starts IEEE 802.11i pre-authentication with AP1. This is an ideal scenario to compare the values obtained from 802.11i pre-authentication with that of network-layer assisted pre-authentication. Both scenarios use RADIUS as AAA protocol (APs implement a RADIUS client). The third scenario takes advantage of network layer assisted link-layer pre-authentication. It involves movement between two APs (e.g., between AP0 and AP1) that belong to two different subnets where 802.11i pre-authentication is not possible. Here, Diameter is used as AAA protocol (PAA implements a Diameter client).

In this third movement scenario, the MN is initially connected to AP0. The MN starts PANA pre-authentication with the PAA which is co-located on the AR in the new candidate target network (nAR in network A) from the current associated network (network B). After authentication, PAA proactively installs two keys, PSKap1 and PSKap2 in both AP1 and AP2 respectively. By doing the key installations proactively, it preempts the process of communicating with AAA server for the keys after the mobile moves to the new network. Finally, because PSKap1 is already installed, AP1 starts immediately the 4-way handshake. We have used measurement tools such as ethereal and kismet to analyze the measurements for the 4-way handshake and PANA authentication. These measurements reflect different operations involved during network-layer pre-authentication.

In our experiment, as part of the discovery phase, we assume that the MN is able to retrieve PAA's IP address and all required information about AP1 and AP2 (e.g. channel, security-related parameters, etc.) at some point before the handover. This avoids the scanning during

link-layer handoff. We have applied this assumption to all three scenarios. Because our focus is on reducing the time spent on authentication part during handoff, we do not discuss the details of how we avoid the scanning.

Types	802.11i Post Authentication		802.11i Pre Authentication		MPA-assisted Layer 2 Preauthentication	
Operation	Non Roaming	Roaming	Non Roaming	Roaming	Non Roaming	Roaming
Tauth	61 ms	599 ms	99 ms	638 ms	177 ms	831 ms
Tconf	--	--	--	--	16 ms	17ms
Tassoc+4way	18 ms	17 ms	16 ms	17 ms	16 ms	17 ms
Total	79 ms	616 ms	115 ms	655 ms	208 ms	865 ms
Time affecting handover	79 ms	616 ms	16 ms	17 ms	15 ms	17 ms

Figure 10: Results of MPA-assisted Layer 2 results

Figure 14 shows the timing (rounded off to the most significant number) associated with some of the handoff operations we have measured in the testbed. We describe each of the timing below. Tauth refers to the execution of EAP-TLS authentication. This time does not distinguish whether this authentication was performed during pre-authentication or a typical post-authentication.

Tconf refers to time spent during PSK generation and installation after EAP authentication is complete. When network-layer pre-authentication is not used, this time is not considered.

Tassoc+4way refers to the time dedicated to the completion of association and the 4-way handshake with the target AP after the handoff.

C.6. Guidelines for handover preparation

In this section, we provide some guidelines for the roaming clients that use pre-authentication mechanisms to reduce the handoff delay. These guidelines can help determine the extent of pre-authentication operation that is needed based on a specific type of movement of the client. IEEE 802.11i and 802.11r take advantage of preauthentication mechanism at layer 2. Thus, many of the guidelines observed for 802.11i-based pre-authentication and 802.11r-based fast roaming could also be applicable to the clients that use MPA-based pre-authentication techniques. However, since MPA operations are not limited to a specific subnet and involve inter-subnet and inter-domain handover the guidelines need to take into account other factors such as movement pattern of the mobile, cell size etc.

The time needed to complete pre-authentication mechanism is an important parameter since the mobile node needs to determine how much ahead of time the mobile needs to start the pre-authentication process so that it can finish the desired operations before the handover to the target network starts. The pre-authentication time will vary depending upon the speed of the mobile (e.g., pedestrian, vs. vehicular) and cell sizes (e.g., WiFi, Cellular). Cell residence time is defined as the average time the mobile stays in the cell before the next handoff takes place. Cell residence time is dependent upon the coverage area and velocity of the mobile. Thus, cell residence time is an important factor in determining the desirable pre-authentication time that a mobile should consider.

Since pre-authentication operation involves six sub-operations as described in Section 7.2 and each sub-operation takes some discrete amount of time, only part of these sub-operations may be completed before handoff depending upon the available delay budget.

For example, a mobile could complete only network discovery and network layer authentication process before the handoff and postpone the rest of the operations to until after the handover is complete. On the other hand if it is a slow moving vehicle and the adjacent cells are sparsely spaced, a mobile could complete all the desired MPA related operations. Finishing all the MPA related operations ahead of time reduces the handoff delay but adds other constraints such as cell residence time.

We give a numerical example here similar to [IEEE-03-084].

D= Coverage diameter,

v= Mobile's velocity,

RTT = round trip time from AP to AAA server including processing time for authentication Tauth

Tpsk = Time spent to install keys proactively on the target APs

If for a given value of D = 100ft, Tpsk = 10 ms, and RTT = 100 ms, a mobile needs to execute only the pre-authentication procedure associated with MPA, then the following can be calculated for a successful MPA procedure before the handoff is complete.

2RTT+Tpsk < D/v

v = 100 ft/(200 ms +10 ms) = ~500 ft/sec

Similarly, for a similar cell size, if the mobile is involved in both pre-authentication and pre-configuration operations as part of the MPA procedure, and it takes an amount of time Tconfig= 190 ms to complete the layer 3 configuration including IP address configuration, then for a successful MPA operation,

2RTT+Tpsk+Tconfig < D/v

v = 100 ft /(200 ms + 10 ms + 190 ms) = ~250 ft/sec

Thus, compared to only pre-authentication part of MPA operation, in order to be able to complete both pre-autentication and pre-configuration operations successfully, either the mobile needs to move at a slower pace or it needs to expedite these operations for this given cell size. Thus, types of MPA operations will be constrained by the velocity of the mobile.

As an alternative if a mobile does complete all the pre-authentication procedure much ahead of time, it uses up the resources accordingly by way of extra IP address, tunnel and extra bandwidth. Thus, there is always a tradeoff between the performance benefit obtained from pre-authentication mechanism and network characteristics, such as movement speed, cell size, and resources utilized.

Authors' Addresses

Ashutosh Dutta
Telcordia Technologies
1 Telcordia Drive
Piscataway, NJ 08854
USA

Phone: +1 732 699 3130
Email: ashutosh.dutta@ieee.org

Victor Fajardo
Telcordia Technologies
1 Telcordia Drive
Piscataway, NJ 08854
USA

Phone:
Email: vf0213@gmail.com

Yoshihiro Ohba
Corporate R&D Center, Toshiba Corporation
1 Komukai-Toshiba-cho, Saiwai-ku
Kawasaki, Kanagawa 212-0001
Japan

Phone:
Email: yoshihiro.ohba@toshiba.co.jp

Kenichi Taniuchi
Toshiba Corporation
2-9 Suehiro-cho
Ome, Tokyo 198-8710
Japan

Phone:
Email: kenichi.taniuchi@toshiba.co.jp

Henning Schulzrinne
Columbia University
Department of Computer Science
450 Computer Science Building
New York, NY 10027
USA

Phone: +1 212 939 7004
Email: hgs@cs.columbia.edu