Chapter 3

Systems analysis of mobility events

In this chapter, I analyze the mobility event resulting out of a mobile node’s handoff. In particular, I analyze the discrete operations of a mobility event in order to design a formal mobility systems model. However, I limit the analysis of the mobility systems model to infrastructure-based mobility only. Infrastructure-based mobility model assumes that the layer 2 and layer 3 network components of the core networks, such as access points, routers, and servers are not mobile and only the last hop of the access network changes. Figure 3.1 shows different functional components in an IP-based infrastructure-based mobility environment. This figure consists of layer 2 and layer 3 points of attachment, configuration agent, authentication agent, authorization agent, and signaling agent that perform the primitive handoff operations. During a mobility event, these network components are engaged in a distributed communication to take care of the handoff related operations. Figure 3.1 also shows how the mobile node changes its point of attachment as it moves between the layer 2 point of attachment, layer 3 point of attachment, different administrative domains and different access technologies. The markings A, B, C, and D in Figure 3.1 show the location of these mobile nodes as the mobile node connects to different point-of-attachment during its movement. The mobile starts from location marked as A and moves to locations marked as B, C, and D. During this process, the mobile is subjected to layer 2, layer 3 and
CHAPTER 3. SYSTEMS ANALYSIS OF MOBILITY EVENTS

administrative domain handoff (e.g., handoff from domain A and domain B). Each of these handoffs (e.g., L2 handoff and L3 handoff) involves different types of handoff operations and different amounts of delays. For example, a layer 2 handoff contributes up to 900 ms delay in 802.11 environment, a layer 3 handoff contributes up to 4 second delay that is inclusive of layer 2 delay and an administrative domain handoff involving heterogeneous access results in a handoff delay up to 18 seconds [DKZ'05]. The networks N1 and N2 are shown as networks with different access technologies (e.g., 802.11 and CDMA). Figure 3.2 shows three different audio output from the mobile when it is subjected to layer 2, layer 3 and heterogeneous handoff. These output waveforms represent the media disruption due to delays contributed by the handoff related operations. Thus, it would be useful to investigate the delays related to each type of handoff operations. In this section, I analyze the delays due to handoff related operations at different layers.

Figure 3.1: Functional components of infrastructure-based mobility
3.1 Introduction

In order to determine the primitive operations that take place during a handoff event, I investigated and analyzed several cellular and IP-based mobility protocols at different layers. I have described the details of handoff operations for a set of cellular and IP-based mobility protocols in Chapter 2. In Table 3.1, I summarize how the various cellular and IP-based mobility protocols perform similar set of operations to support a handoff event. In particular, I illustrate how these mobility protocols carry out different handoff related functions, namely network discovery, resource discovery, network detection, configuration, authentication, encryption, binding update and media rerouting. The column headings in the table represent the primitive operations that are part of a handoff event for different types of mobility protocols that are represented in the row headings.

Operational aspects of these functions play an important role in the design of any specific mobility protocol. For example, the operational aspect of a discovery function could include many factors like, number of messages between the mobile and the server, discovery mechanism (e.g., network layer or application layer). Performance of any mobility protocol depends upon how efficiently these functions operate. In order to get a better understanding, I have performed a comparative analysis of application layer mobility protocol (e.g., SIP-based mobility) and network layer mobility protocol (e.g., MIPv6) and have
<table>
<thead>
<tr>
<th>Mobility Type</th>
<th>Access Type</th>
<th>Network Discovery</th>
<th>Resource Discovery</th>
<th>Triggering Technique</th>
<th>Detection Technique</th>
<th>Configuration</th>
<th>Authentication</th>
<th>Encryption</th>
<th>Binding Update</th>
<th>Media</th>
<th>Relocating</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSM</td>
<td>TDMA</td>
<td>BCCH</td>
<td>FCCH</td>
<td>Channel Strength</td>
<td>SCH</td>
<td>TMSI</td>
<td>SRES A3</td>
<td>DES</td>
<td>MSC Control</td>
<td>Anchor</td>
<td></td>
</tr>
<tr>
<td>WCDMA</td>
<td>CDMA</td>
<td>PILOT</td>
<td>SYNC channel</td>
<td>Channel strength</td>
<td>Frequency</td>
<td>TMSI</td>
<td>SRES/A3</td>
<td>AES</td>
<td>Network Control</td>
<td>Anchor</td>
<td></td>
</tr>
<tr>
<td>IS-95</td>
<td>CDMA</td>
<td>PILOT</td>
<td>SYNC channel</td>
<td>Channel strength</td>
<td>RTC</td>
<td>TMSI</td>
<td>Diffie-Hellman CAVE</td>
<td>AES</td>
<td>MSC</td>
<td>PDSN</td>
<td>MSC</td>
</tr>
<tr>
<td>802.11</td>
<td>CSMA/CA</td>
<td>Beacon 11R</td>
<td>SNR at Mobile</td>
<td>Scanning Channel number SSID</td>
<td>802.1X EAPcL</td>
<td>WEP WPA 802.11i</td>
<td>Associate</td>
<td>IAPP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cellular IP</td>
<td>Any</td>
<td>Gateway beacon</td>
<td>AP beacon</td>
<td>AP beacon ID</td>
<td>GW Beacon</td>
<td>MAC Address</td>
<td>IKE</td>
<td>AES</td>
<td>Route Update</td>
<td>Intermediate router</td>
<td></td>
</tr>
<tr>
<td>MIPv4</td>
<td>Any</td>
<td>ICMP RA, FA Advertising</td>
<td>ICMP</td>
<td>FA advertisement, L2 assn.</td>
<td>FA-CoA CoCoA</td>
<td>IKE PANA EAP</td>
<td>AES</td>
<td>MIP Registration</td>
<td>FA</td>
<td>RFA HA</td>
<td></td>
</tr>
<tr>
<td>MIPv6</td>
<td>Any</td>
<td>Stateless RA Proactive</td>
<td>CARD 802.21 802.11 R</td>
<td>Router Adv.</td>
<td>Router Prefix</td>
<td>CoA</td>
<td>IKE PANA EAP</td>
<td>AES</td>
<td>MIP Update MIP</td>
<td>CH</td>
<td>MAP HA</td>
</tr>
<tr>
<td>SIP-based mobility</td>
<td>Any</td>
<td>Stateless RA ICMP Router</td>
<td>802.21 802.11R</td>
<td>L3 Router Adv.</td>
<td>Router Prefix ICMP</td>
<td>CoA AOR Register</td>
<td>IKE EAP Invite Exchange</td>
<td>AES</td>
<td>Re-INVITE</td>
<td>CH</td>
<td>RTP-Trans</td>
</tr>
</tbody>
</table>
illustrated how each of these primitive operations is carried out for both of these mobility protocols [DSC+06]. I briefly describe how some of the handoff operations are taken care of by each of these two mobility protocols.

### 3.1.1 Comparative analysis of network layer and application mobility protocol

In Chapter 2, I have described mobility management techniques for both SIP-based mobility and mobile IPv6 (MIPv6). However, in this section, I compare how some of the basic operations such as discovery, registration, binding update, configuration, location management, tunneling, and security operations are done for each of these two mobility protocols.

In SIP-based mobility management, as part of the registration process, the mobile updates its IP address with the visited SIP proxy or home proxy. In case of MIPv6, the mobile sends the binding update to the home agent and corresponding host. Thus binding update to the home agent can be regarded as the registration process for MIPv6.

Both MIPv6 and SIP-based mobility do not require any foreign agent in the network and thus use co-located care-of-address as the new identifier. SIP-based mobility for IPv4 networks mostly depends upon DHCP for IP address configuration. Both SIP-based mobility for IPv6 networks and MIPv6 however use stateless auto-configuration to configure a layer 3 identifier.

In the absence of route optimization, mobile IPv6 tunnels payload packets between the mobile node and the home agent in both directions. This specific tunneling mechanism uses IPv6 encapsulation as specified by RFC 2473 [CD98]. In addition to the extra headers assigned to the original packet, additional time is spent due to processing of encapsulation and de-capsulation. SIP-based mobility on the other hand does not make use of tunneling as the media travels directly between the CH and MH. Thus, processing delay due to encapsulation and tunneling overhead is avoided when SIP-based mobility is used.

Binding update has been defined in Chapter 2. In IP-based environment, it is the process
of notifying the correspondent node or other networking node such as home agent about the new layer 3 identifier of the mobile so that the data can get forwarded to the new address of the mobile after the handoff. In case of Mobile IPv6, as the mobile obtains a new care-of-address either via stateful DHCP server or stateless auto-configuration it notifies the correspondent host and the home agent. Since route optimization is an inbuilt mechanism for MIPv6, the new data during mid-session mobility does not need to get rerouted via home agent. On the other hand, SIP-based mobility sends the binding update as part of its Re-INVITE signal to the correspondent host only since there is no home agent for SIP-based mobility.

Mobile IPv6 by itself does not provide fast-handoff mechanism. However, there are extensions to Mobile IPv6, namely, FMIPv6 [Koo05] and HMIPv6 [SCeMB06] that provide fast-handoff during mid-session mobility. FMIPv6 provides fast-handoff mechanism by introducing several reactive and proactive mechanisms, whereas HMIPv6 introduces a mobility anchor point (MAP) to take care of intra-domain mobility. Similarly, SIP-based mobility has been extended to provide fast-handoff by introducing a network entity called B2BUA (Back-to-Back-User Agent) in the hierarchy to limit the binding update or by forwarding the transient traffic from the previous network.

Both of these mobility protocols provide security for the signaling and data. MIPv6 takes advantage of network layer security such as IPsec to protect the signaling between the mobile and home agent. It can also use IPsec tunnel instead of IP-IP tunnel between the mobile and home agent to carry the tunneled data. SIP-based mobility on the other hand can provide a multilayer security. It can either choose IPsec to provide security at layer 3 for both signaling and media (RTP) or it can use S/MIME to secure the SIP signaling and use secure RTP (SRTP) to secure the media stream. However, SRTP can also be used to protect the data in case of MIPv6, but it will need a separate key distribution architecture unlike SIP-based mobility where the key is distributed by Invite exchange method.

Mobile IPv6 does provide return routability support by using CTI (Care-of-Test Init)
and HTI (Home Test Init) messages. This actually verifies the new care-of-address of the mobile before the binding update is sent out. While it helps to avoid session hijacking etc., it does add delay to the binding update procedure. SIP, on the other hand, does not support inherent return routability testing, but new care-of-address of the mobile can be verified by using cryptographic technique such as SIP identity [PJ06].

### 3.2 Analysis of handoff components

Based on the analysis of the basic operations associated with both IP-based mobility protocols and cellular protocols, I categorize the handoff process into six main phases, namely, network discovery and selection, network attachment, configuration, security association, binding update and media reroute. Figure 3.2 decomposes the mobility event into several processes and sub-processes.

![Diagram of Systems decomposition of handoff](image)

**Figure 3.3: Systems decomposition of handoff**

Each of these operations involves several network elements, namely mobile node, access point, router, server and correspondent node. Figure 3.3 shows how each of these
CHAPTER 3. SYSTEMS ANALYSIS OF MOBILITY EVENTS

sub-operations involves several components in the network at several layers. For example, a layer 2 discovery involves interaction between the mobile node and the access point, whereas layer 3 discovery involves mobile node, access point and the router in the network. On the other hand, a binding update involves the core components in the network, such as home agent, server and core routers in the network. I describe below these handoff operations in details.

![Diagram of handoff functions distributed across network elements](image)

Figure 3.4: Handoff functions distributed across network elements

Handover initiation or handover preparation precedes a handover process. Handover initiation process does not contribute to the handover delay. Depending upon who initiates the handover and who has the primary control over the handover, a handoff process can be either mobile-initiated handover, network-initiated handover, mobile-controlled handover or network-controlled handover [EE04]. During this phase, the mobile or the network node determines the need for the handoff based on the measurement on the mobile, and initiates the preparation for the handoff operation. For example, during a mobile-initiated handoff, based on Signal-to-Noise Ratio (SNR) or the quality of service of the media traffic, the mobile makes the decision about the impending handoff and starts the network discovery
process to determine the best available network to connect to. A handover process may result in re-initiation of layer 2 or layer 3 operations based on whether MN loses the layer 2 or layer 3 connectivity, respectively. For example, in case the mobile does not receive the layer 3 router advertisement, it initiates a layer 3 handoff process without necessarily disturbing the layer 2 association. On the other hand, a layer 2 unreachability event resulting out of low signal-to-noise ratio or non-receipt of beacon will lead to handover initiation process in layer 2 that may or may not involve handoff at layer 3.

The IEEE 802.21 [802b] working group has defined several event service primitives as part of its Media Independent Handover Function (MIHF). “Link Up”, “Link Going Down”, and “Link Down” are some examples of such primitives that can be used to provide the status of lower layer events and expedite the handoff process. For example, since the layer 2 association takes place before any other upper layer operations, it helps to send a layer 2 event notification, such as “Link up” to execute the upper layer mobility function like layer 3 configuration. Signal-to-Noise Ratio (SNR) threshold can possibly generate such an event notification. Thus, an event notification from lower layer helps to establish a successful link in the new point of attachment in an expedited manner.

### 3.2.1 Network discovery and selection

The very first phase of the handoff process is network discovery and selection phase. During this phase, the mobile or network discovers possible new points of attachment. This phase involves discovering both the neighboring networks and the required resources for handoff within the network. Once the target network is discovered, several resource parameters within the target network are retrieved including channel number, bandwidth, encryption algorithm, authentication server, registration server, and configuration server. The resource discovery process helps the mobile to associate with proper channel number and proper authentication parameters in the new network so that it can communicate successfully. Jari et al. [AAKB08] provide an overview of network discovery and selection, and its
applicability to handoff. Based on the type of access technology, the discovery process can be passive, where the node listens for network announcements, or active where the node solicits network announcements. Different types of discovery process take place for each type of network, such as discovery of a new Location Area (LA) for GSM or discovery of new Routing Area (RA) in case of GPRS. For IP-based networks, discovery process can span across all the layers and include layer 2 PoA, layer 3 PoA, subnet or domain discovery. Here a domain is defined as an administrative domain.

Each layer 2 access technology provides different means of out-of-band mechanism for discovering the networks and resources. For example, GSM uses BCCH (Broadcast Control Channel), CDMA uses pilot channel and 802.11 uses active scanning and passive scanning to discover the new point-of-attachment. Based on the type of access characteristics of the target networks, discovery of the appropriate network takes different amounts of time. For example, in IEEE 802.11 network, without any optimization, active scanning operation takes about 500 ms [MSA03] including the channel probe delay, authentication and association. Network discovery timing in GSM network is determined by the frequency of broadcast channel. In case of IP-based mobility, some of the upper layer detection mechanisms, such as foreign agent advertisement or router advertisement can help discover the servers in the network.

Network selection is a process by which a mobile node or a network element analyzes the information discovered about its neighboring networks, and then selects a network to connect to. The selection may be based on criteria such as required QoS, cost or user preferences. An appropriate selection mechanism helps the overall resource optimization process and increases the probability of successful handoff to the target network.

### 3.2.2 Network attachment

After the mobile has selected the target network, it attempts to connect to the new network point of attachment. Network attachment can take place in several layers. A mobile at-
taches to an 802.11 access point by means of layer 2 association and connects to a router by means of layer 3 association. Event notification from the lower layers, such as the availability of the new point of attachment or sudden lapse of an existing connection is usually passed to the upper layers in order to initiate the subsequent handoff related functions in an expedited manner. The IETF is currently working on standardizing a protocol for detection of network attachment called DNA [Cho05] that involves mechanisms both at layer 2 and layer 3 and can notify the upper layer about the network detection.

3.2.3 Configuration

Configuration phase helps to prepare the re-connection path of the mobile. During this phase, the mobile obtains a temporary identifier in the network the mobile is visiting, discovers servers such as outbound SIP server and default gateway, and maps the address with the appropriate network entity in the network. Configuration phase can be split into the following sub-phases, namely identifier acquisition, duplicate address detection and address resolution. During identifier acquisition process, the mobile acquires a new temporary L2 or L3 connection identifier at the new point of attachment of the network. For example, it could include Care-of-Address (CoA) in case of IP-based mobility or TMSI (Temporary Mobile Subscriber Identity) in case of GSM. In IPv4-based networks, the mobile uses DHCP server to obtain the care-of-address, whereas in IPv6-based network the mobile obtains the address either from a DHCP server or using stateless auto-configuration method.

Duplicate address detection process tests the uniqueness of the mobile’s address in the network. Address uniqueness testing is performed differently by IPv4 and IPv6 networks. For example, when an MN uses stateless address auto-configuration to assign an IPv6 address to its interface after getting a new RA (Router Advertisement), it sends a neighbor solicitation on the local link in order to verify whether any other node on the link has the same address. When the pre-determined time elapses and the MN does not receive any
reply, the MN assumes that no other node on the link has this address and finally assigns that address to its interface. IPv4 node uses ARP broadcast to determine if any other node in the network has the same address.

Once the address uniqueness has been determined, the mobile and the default router keep the mapping between the IP address and MAC address of each other using address resolution mechanism. This address resolution mechanism helps the mobile and the default router to communicate with each other at layer 2 in the same subnetwork.

In general, completion of these processes requires a series of signaling messages between the mobile and servers in the network and thus contributes to the handoff delay.

### 3.2.4 Security association

A security association is defined as a secure channel between two endpoints that applies a security policy and keys to protect the information. Before a new communication path is established between the end-points, the mobile node needs to authenticate itself and then establish security association with other network nodes, such as routers, access points or the correspondent host. Establishing a security association involves authentication and authorization and key derivation procedures. The authentication and authorization processes allow the mobile to get access to the network resources. An authentication process includes exchange of messages between the mobile, authenticator and authenticator server e.g., AAA server in the network. Successful authentication process generates the key that can be used to encrypt the data. Encryption process itself is considered to be separate from authentication and does not affect the handoff delay, unless the type of encryption algorithm changes after the handoff. However, the process of deriving the encryption key contributes to the handoff delay.

Security association may take place at several layers of the protocol stack. For example, during a layer 3 movement in an 802.11 access network, EAP messages need to be carried between the mobile node, authenticator and authentication server to generate the shared
master key followed by a 4-way handshake between the mobile node and the access point to generate the PTK (Pairwise Transient Key) that can be used for encryption at layer 2. In an IP-based network, ISAKMP (Internet Security Association and Key Management Protocol) [MSST98] defines the mechanism for establishing the security association at layer 3 using IPSec. An IPSec security association is defined by the encryption algorithm, the authentication algorithm, and the shared session key.

Each mobility protocol also uses different mechanism for authentication. As shown in Table 3.1, the mobile uses SRES and A3 algorithm for the authentication in GSM. For 802.11 access networks, the mobile can use open system authentication or shared-key authentication WEP or IEEE 802.11i-based authentication in layer 2. Fathi et al. [FKC+05] demonstrate how different authentication mechanisms affect the association and transport delay at layer 2. Similarly, layer 2 access independent network layer authentication protocol such as PANA (Protocol for carrying Authentication to Network Access) [JLO08] add delay during layer 3 authentication. Georgiades [Geo04] shows that it takes up to 4 seconds to complete the authentication and authorization process. These operations could use a combination of EAP (Extensible Authentication Protocol) [ABV+04] over layer 2 for IEEE 802.1x-based authentication and EAP-TLS (Transport layer Security) [AS99] for layer 3 authentication. At the time of reconnection after handoff, re-authentication process adds to the handoff delay. An authentication process is followed by authorization process in some cases. For example, an 802.1x-based or PANA-based authenticator usually communicates with an AAA server to complete the authorization process. In case of inter-domain mobility, the AAA server changes. Thus, the authentication process is further delayed as the mobile needs to go through the authorization process using the new AAA server.

3.2.5 Binding update

Binding update is the process by which a mobile can update its newly obtained network identifier so that the data after handoff can be rerouted to the new destination. A binding
update process consists of three main phases, namely identifier update, identifier mapping and binding cache update. This identifier update process associates the new network identifier with the permanent identifier of the mobile. As the mobile connects to a new point of attachment and obtains a new temporary network identifier (e.g., TMSI in GSM, COA in MIPv6, FA-COA in MIPv4) in the new network, it needs to update the correspondent host and home agent so that the packets can be routed to the new destination.

Until the re-association of the new identifier is complete and binding cache in the correspondent node or home agent is updated, the transient in-flight data continue to go to the old network and is lost in the absence of any optimization mechanism, such as buffering or packet forwarding. For application layer mobility, binding update process also includes registration process. By means of registration process, the mobile establishes the mapping between the permanent locator, e.g., URI (Universal Resource Identifier) the temporary identifier, e.g., care-of-address for proper location management function. An optimized or hierarchical registration process expedites locating a user and provides faster delivery of the new data.

Binding update also needs to be authenticated. For example, MIPv6 introduces return routability procedure and adds two additional messages, namely CTI (Care-of-Test Init) and HTI (Home-Test Init) to obtain the binding key that can authenticate the binding update message. However, this process contributes to additional delay for the binding update procedure after every handoff. Binding update can also be secured by way of IPSec. The security association for binding update in an IP-based network can be uniquely identified by a tuple consisting of a Security Parameter Index (SPI), an IP destination address, and a security protocol (AH or ESP) identifier.
3.2.6 Media re-routing

A media rerouting phase is the last phase during the handover process and follows the binding update process. This phase involves re-routing of the media so that the delivery of data changes from the old path to the new path according to the pre-defined service guarantee.

Once the binding update is complete, the home agent or correspondent host updates its binding cache and the data from the correspondent node gets routed to the mobile’s new location. Media delivery can take place in several ways. In one way, the media delivery can take place using the direct path between the CN and the MN. In second scenario, the media is delivered using an indirect path and uses a network entity, namely home agent. However, until the time the new data is sent directly to the mobile’s new point of attachment (nPnA), there are packets in-flight to the previous point-of-attachment. Media re-routing does also take care of re-routing in-flight packets. In both the cases, the in-flight data can be captured and redirected to the new point of attachment. Media rerouting process may include several elementary operations, such as encapsulation, decapsulation, tunneling, buffering, and store-and-forward techniques. During the media re-routing process, transient data may get lost or may get delayed because of these operations. There is data overhead associated with encapsulation, decapsulation and tunneling operations. Thus, there is a need to optimize these operations to ensure that the media delivery delay is reduced after the handoff. Optimization techniques for the media delivery are often defined as route optimization methodologies. I describe some of these route optimization techniques in Chapter 5. As an example, operations in the network, such as buffering or forwarding mechanisms help to reduce packet loss but add delay to the packet traversal. Thus, the buffering period needs to be adjusted to compromise between packet loss and one-way packet delay.
3.3 Effect of handoff across layers

In this section, I illustrate how these basic handoff operations affect multiple layers in an IP-based network. Based on the type of mobility, a subset of operations are executed at each layer and the overall handoff delay is contributed by delays at all layers during a mobility event. For example, during an intra-domain movement, the mobile is not subjected to the delay due to authorization process unlike inter-domain mobility. Similarly, a layer 2 handoff does not involve delays due to other layer 3 operations, namely layer 3 identifier acquisition or duplicate address detection. Following is a description of how optimization techniques can be useful to each of these basic operations at every layer.

3.3.1 Layer 2 delay

In 802.11 environment, channel scanning, probing, authentication, and association are the basic functions that contribute to the delay before a mobile completes the network attachment at layer 2. Scanning is considered to be a discovery process in layer 2. Encryption and user authentication using WPA (Wi-Fi Protected Access) in conjunction with 802.1X [802a] and EAP (Extensible Authentication Protocol) contribute to the additional delays because of the associated 4-way handshake between the mobile and the access point. Mishra et al. [MSA03] provide a comprehensive delay analysis of the basic operations associated with a layer 2 handoff. Shin et al. [SSFR04] also discuss the probe delays, authentication and association delays during layer 2 handoff in 802.11 networks. Both of these studies demonstrate that the probe delay constitute ninety percent of the total layer 2 handoff delay.

I have also taken measurements in the experimental testbed that I have created to study the layer 2 handoff related operations for two different operating systems, namely Linux and Windows, and used different layer 2 drivers, namely Aironet, Orinoco, DLink, and Centrino. From a thorough analysis of the layer 2 handoff event I found that scanning
and probing operations contribute to most of the delay. For example, in our experimental setup, using active scanning with Orinoco driver in a Linux environment, it takes almost 100 ms for probing action to complete. This is followed by layer 2 open authentication and association that take about 2 ms and 20 ms, respectively. I have also experimented with layer 2 authentication, such as IEEE 802.11i and have found average EAP delay and 4-way handshake delay to be 79 ms and 616 ms for non-roaming and roaming cases, respectively. Roaming scenario involves interaction with AAA server during mobile's handoff, where as authentication is limited to the local authentication agent for non-roaming scenario.

3.3.2 Layer 3 delay

In an IP-based environment, network association at layer 3 involves several basic operations, such as discovery of layer 3 point of attachment (e.g., default gateway discovery), IP address acquisition, duplicate address detection, neighbor reachability, local access authentication, and authorization. Each of these operations involves a number of message exchanges between the mobile and other network entity such as router, DHCP server, authentication server. From the periodic router advertisement, the mobile can discover the new layer 3 point of attachment. There are several protocols, such as DHCP [Dro97], DHCPv6 [DEB+03], PPP [Sim94] and stateless auto-configuration [TN98] that help the mobile to acquire IP address. I have measured the time taken by each of these IP address configuration methods and have explained them in [DMCS06]. Table 3.2 shows IP address acquisition delays due to different types of configuration protocols including zeroconf [CAG05], static (manual configuration) and proactive IP address acquisition technique. Time taken to configure an IP address using each of these protocols differs due to various factors, such as variation in the number of signaling messages to acquire the IP address and duplicate address detection technique. In a typical inter-domain mobility scenario, there are additional operations, such as re-authentication, re-authorization and profile verification by the AAA servers during the handoff process. In some cases, the AAA-related messages traverse all
the way to the home AAA server before the network service is granted to the mobile in the new network. Thus, it is desirable to have an optimized method of interaction between the mobile and AAA server during the inter-domain handoff. After a layer 3 identifier is re-configured, other layer 3 functions, such as the binding update from the mobile and media redirection at the correspondent host contribute to the additional delay.

Table 3.2: IP address acquisition delay

<table>
<thead>
<tr>
<th>Layer 3 configuration methods</th>
<th>DHCP w/ ARP</th>
<th>DHCP w/o ARP</th>
<th>IPv6 stateless</th>
<th>DHCPv6</th>
<th>PPP</th>
<th>FA CoA</th>
<th>Zero Conf</th>
<th>Manual</th>
<th>Proactive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address acquisition delay</td>
<td>4 s</td>
<td>400 ms</td>
<td>160 ms</td>
<td>500 ms</td>
<td>8 s</td>
<td>2 s</td>
<td>5 s</td>
<td>100 ms</td>
<td>4 ms</td>
</tr>
</tbody>
</table>

### 3.3.3 Application layer delay

A mobile node suffers from application layer delay during a handoff operation, when some of the handoff related operations are performed at the application layer. Application layer delay is mostly attributed due to the operations such as binding update delay at the application layer, processing delay at the end hosts, registration, upper layer encryption, such as TLS (Transport Layer Security), and SRTP (Secured RTP). SIP-based mobility is a typical example where binding update is done at the application layer. SIP-based binding update (e.g., Re-INVITE) contributes to the delay because of three-way-exchange between the mobile and the correspondent host during identifier rebinding process.

### 3.3.4 Handoff operations across layers

Figure 3.5 illustrates a sample protocol flow where the mobile is subjected to subnet handoff and uses MIPv6 as the mobility protocol.

In this specific example, the mobile moves from old layer 2 point of attachment (oPoA) to a new layer 2 point of attachment (nPoA) and in the process it disconnects from previous Access Router (pAR) and connects to the new Access Router (nAR) resulting in change
in layer 3 point of attachment. This figure shows the message exchange between different components in the network that are used to take care of handoff operations, namely layer 2 discovery, layer 3 discovery, authentication, layer 2 security association, address acquisition, duplicate address detection and binding update. Each of these operations needs different number of signaling messages and carry varying amounts of data payload. Thus, consumption of systems and network resources (e.g., cpu cycles, battery power, bandwidth) and amount of time taken to complete each of these operations will vary.

In order to study the effect of different components on handoff delay, I experimented with secured inter-domain mobility using both network layer (e.g., Mobile IP) and application layer (e.g., SIP-based) mobility techniques [DAC*03] and [DDL*04]. Inter-domain mobility introduces additional operations, such as local authentication and interaction with an authorization server resulting in additional delay. These mobility prototypes include a
Figure 3.6: Call flow for mobile IP-based inter-domain handoff

combination of network detection, registration, configuration, authentication, security, location management functions, and roaming support for the wireless Internet [DVC+01]. I have used DRCP, a variant of DHCP for configuration, PANA [JLO08] for secure access control, Diameter [CLG+03] for authorization, and ESP (Encapsulating Security Payload) within IPSec [KA98a] and SRTP (Secured RTP) [BMN+04] for encryption.

Figures 3.6 and 3.7 show the call flows for mobile IP-based and SIP-based inter-domain handoff, respectively. ERC1 and ERC2 are the two edge router controllers that also act as DRCP server, PANA server and IPSec end point. These flows demonstrate the protocol interaction among several functional components, such as SIP user agent or Mobile IP client, DRCP server, PANA server, home agent, IPSec server and AAA server when the mobile moves from home domain to visited domain and return to the home domain. During this movement, the mobile changes its layer 2 point-of-attachment, layer 3 point-of-attachment, and then re-establishes layer 3 authentication, authorization and security association using IPSec. During the inter-domain handoff the mobile needs to communicate with the visited
AAA server for re-authorization.

Table 3.3 shows the timing delay for some of the functional components during inter-domain handoff when SIP-based mobility and MIP are used as the mobility protocols. End systems processing time are not included in this table. These experimental results demonstrate delays due to several handoff components, namely layer 2 beacon interval, subnet, and domain discovery, IP address acquisition, local authentication, authorization, and delay due to binding update, such as SIP re-INVITE and MIP registration. Pre-handoff media represents the time when the media is received prior to handoff in the old network and post-handoff media represents the time media is received in the new point of attachment after the handoff. For both the cases, IKE (Internet Key Exchange) took almost 5 seconds for successful establishment of the IPSec-based security association.

Table 3.4 summarizes how the basic operations of a mobility event as described in Section 3.1 are performed across different layers in an IP-based environment with 802.11 as the access media. Depending upon the layer at which the handoff takes place, (e.g.,
Table 3.3: Experimental timing for handoff components

<table>
<thead>
<tr>
<th>Mobility Type</th>
<th>Pre-handoff media</th>
<th>Layer 2 beacon adv.</th>
<th>Router advertisement</th>
<th>Layer 3 configuration</th>
<th>Layer 3 authentication delay</th>
<th>IKE process</th>
<th>Binding update</th>
<th>Post-handoff media</th>
<th>In-handoff delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIP</td>
<td>51.7 s</td>
<td>120 ms</td>
<td>500 ms</td>
<td>80 ms</td>
<td>10 ms</td>
<td>5130 ms</td>
<td>240 ms</td>
<td>53 s</td>
<td>1.3 s</td>
</tr>
<tr>
<td>MIP</td>
<td>23.8 s</td>
<td>120 ms</td>
<td>500 ms</td>
<td>80 ms</td>
<td>10 ms</td>
<td>4580 ms</td>
<td>20 ms</td>
<td>31.1 s</td>
<td>7.3 s</td>
</tr>
</tbody>
</table>

layer 2 or layer 3) appropriate operations at those layers are performed. For example, if a mobile's handoff operation does not involve change in subnet then many of the layer 3 related operations are not performed during the mobility event.

Table 3.4: Mapping of basic handoff operations across layers

<table>
<thead>
<tr>
<th>Layers</th>
<th>Basic handoff operations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Discovery</td>
</tr>
<tr>
<td>Layer 2</td>
<td>Scanning</td>
</tr>
<tr>
<td>Layer 3</td>
<td>Router advertisement</td>
</tr>
<tr>
<td>Application Layer</td>
<td>AAA domain discovery</td>
</tr>
</tbody>
</table>

While optimization of each of these specific operations minimizes overall handoff delay, scheduling among different tasks across layers can also lead to variety of optimization techniques. For example, it is not necessary for all the layer 2 operations to complete before layer 3 operations can start. Scheduling a layer 3 discovery process during layer 2 discovery process or configuring a layer 3 identifier before a layer 2 identifier is configured will help minimize the handoff delay. Figures 3.8, 3.9 and 3.10 illustrate the effect of non-
optimized, reactive and proactive handoff on the mobile, respectively. I used Mobile IPv6 for non-optimized handoff and used Fast Mobile IPv6 to obtain the results for both reactive and proactive handoffs. As shown in these figures, in this specific laboratory testbed environment, a mobile suffers from almost 7 seconds delay and lose about 70 packets due to handoff when MIPv6 was used without any optimization. Whereas the reactive optimization associated with Fast MIPv6 [Koo05] reduces the delay to almost 1.5 seconds. Finally, when FMIPv6’s proactive optimization techniques are applied, the handoff delay is little less than handoff delay during reactive optimization but the packet loss is reduced to zero due to buffering techniques applied at the next access router. It is important to note that these experiments did not include any optimization at layer 2.

I describe my proposed optimization techniques for each of the handoff components in Chapter 5.
Figure 3.9: Results of FMIPv6 reactive handoff

Figure 3.10: Results of FMIPv6 proactive handoff