Chapter 2

Introduction to mobility management for multimedia

Mobility management consists of two components: location management and handoff management. Location management enables the network to discover the current point of attachment of the mobile user so that the new connection can be established when a new multimedia call arrives. Handoff management, often known as terminal mobility, allows the network to maintain the user’s connection binding as the mobile node moves from one attachment point to another in the network. I focus on handoff management in my thesis.

As a mobile goes through a handover process, it is subjected to connection disruption because of the rebinding of its association at several layers of the protocol stack. Delays incurred due to rebinding within each of these layers affect the ongoing multimedia application and data traffic within the client. Several basic operations are associated with the re-establishment of the binding process across these layers. These operations can be affected by several factors, such as access characteristics (e.g., bandwidth, channel characteristics), access mechanism (e.g., CDMA, CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance), TDMA (Time Division Multiple Access)), re-configuration of identifiers, re-authentication, re-authorization, and rebinding of security associations at all
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

layers.

Mobility protocols have evolved over a period of last three decades. Based on access characteristics and bandwidth, these can be classified into five main categories: 1G cellular, 2G cellular, 3G cellular, 4G, and IP-based mobility. The definition section in Appendix C define 1G, 2G, 3G and 4G. These mobility protocols exhibit certain similar functionalities during their handoff operation. I highlight these similarities when I describe these protocols later in the chapter. Figure 2.1 shows the evolution of access technologies for the generation of mobility standards. Table 2.1 shows the details of the access characteristics, frequency and data rate of these protocols. In this section, I briefly discuss how mobility operation is performed in 1G, 2G, 3G, and 4G access networks including IP-based mobility protocols and but analyze the associated abstract functions in Chapter 3.

Figure 2.1: Evolution of wireless access technologies
## Table 2.1: Access characteristics of cellular protocols

<table>
<thead>
<tr>
<th>Generation</th>
<th>System</th>
<th>Channel spacing</th>
<th>Access type</th>
<th>Uplink data rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1G</td>
<td>AMPS</td>
<td>30 kHz</td>
<td>FDMA</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>TACS</td>
<td>25 kHz</td>
<td>FDMA</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>NMT</td>
<td>25 kHz</td>
<td>FDMA</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>NTT</td>
<td>25 kHz</td>
<td>FDMA</td>
<td>N/A</td>
</tr>
<tr>
<td>2G</td>
<td>GSM</td>
<td>200 kHz</td>
<td>TDMA</td>
<td>9.6 kb/s</td>
</tr>
<tr>
<td></td>
<td>PDC</td>
<td>30 kHz</td>
<td>TDMA</td>
<td>42 kb/s</td>
</tr>
<tr>
<td></td>
<td>IS-136</td>
<td>30 kHz</td>
<td>F/TDMA</td>
<td>48 kb/s</td>
</tr>
<tr>
<td></td>
<td>IS-95 (A)</td>
<td>1.25 MHz</td>
<td>F/CDMA</td>
<td>14.4 kb/s</td>
</tr>
<tr>
<td></td>
<td>iDEN</td>
<td>25 kHz</td>
<td>F/TDMA</td>
<td>24 kb/s</td>
</tr>
<tr>
<td>2.5G</td>
<td>GPRS</td>
<td>200 kHz</td>
<td>TDMA</td>
<td>45 kb/s</td>
</tr>
<tr>
<td></td>
<td>EDGE</td>
<td>200 kHz</td>
<td>TDMA</td>
<td>236 kb/s</td>
</tr>
<tr>
<td></td>
<td>IS-95 (B)</td>
<td>1.25 MHz</td>
<td>F/CDMA</td>
<td>115 kb/s</td>
</tr>
<tr>
<td>3G</td>
<td>CDMA2000 1X UMTS/WCDMA</td>
<td>1.25 MHz</td>
<td>CDMA/CDMA/TDMA</td>
<td>144 kb/s / 2 Mb/s</td>
</tr>
<tr>
<td></td>
<td>CDMA2000 1xEV-DO</td>
<td>1.25 MHz</td>
<td>CDMA</td>
<td>2 Mb/s</td>
</tr>
<tr>
<td>4G</td>
<td>LTE</td>
<td>20 MHz</td>
<td>OFDMA</td>
<td>50 Mb/s</td>
</tr>
<tr>
<td></td>
<td>WiMAX</td>
<td>2.5 GHz</td>
<td>OFDM</td>
<td>40 Mb/s</td>
</tr>
<tr>
<td></td>
<td>UMB</td>
<td>5 MHz</td>
<td>OFDMA</td>
<td>75 Mb/s</td>
</tr>
</tbody>
</table>
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

2.1 Cellular 1G

1G refers to the first generation of wireless telephone technology. 1G cellular is based on analog technology and was introduced in the 1980s and continued until being replaced by 2G digital telecommunications. Several 1G cellular standards were developed in different countries. One such standard is NMT (Nordic Mobile Telephone) used in Nordic countries, Eastern Europe and Russia. Others include AMPS (Advanced Mobile Phone System) used in the United States, TACS (Total Access Communications System) in the United Kingdom, JTACS in Japan, C-Netz in West Germany, Radiocom 2000 in France, and RTMI (Radio Telefono Mobile Integrato) in Italy. In 1979, the first analog cellular system, the Nippon Telephone and Telegraph (NTT) system was started. In 1981, Ericsson Radio Systems AB deployed the Nordic Mobile Telephone (NMT) 900 systems and in 1983 and ATT deployed the Advanced Mobile Phone Service (AMPS) as a trial in Chicago. However, the two most important 1G systems deployed in the world are AMPS and TACS. All the 1G systems use Frequency Division Multiple Access but each system works on different frequency range. For example, AMPS operates on 800 MHz frequency band, whereas TACS and NMT450 operate on 900-MHz and 450-MHz frequency bands, respectively. I briefly introduce the 1G architecture and describe how handoff is performed in 1G network.

2.1.1 Systems architecture

Figure 2.2 shows the simple system architecture of 1G system. The main components of the system are mobile host (MH), base station (BS), base station controller (BSC), and mobile switching center (MSC) or often known as mobile telephone switching office (MTSO) in AMPS. Base stations are considered to be the point of attachment (PoA) for the mobile host in a specific radio cell. MSC acts like a mobility anchor agent in the network. Definition of these terms are given in the glossary as part of Appendix B.

Mobile is assigned a mobile identifier number (MIN) that is equivalent to its home ad-
dress. It includes an area code identifying the home address area, a three digit exchange number and a four digit subscriber identification number. Each mobile also has an electronic serial number (ESN) that is permanently assigned by the manufacturer. This is equivalent to a device identifier and helps to secure the mobile. While the voice traffic is transmitted using a traffic channel, signaling among the network components are done by control channels that are used to initiate the calls and for handoff related operations. There are two kinds of control channels: forward and reverse. For AMPS, the forward channel is known as the FOrward Control Channel (FOCC) and the other one is REverse Control Channel (RECC). FOCC is transmitted continuously by the base station so that it can be received by all the mobiles. The mobile uses FOCC to associate with the network and perform the handover. Mobile uses RECC to register with the network.

The network and base stations are provided with identification codes. For example, in USA, the System Identifier (SID) is assigned by FCC along with the carrier frequencies for each geographical area. Mobile host uses the system identifier to ensure that it is on the
correct network. The MSCs are linked together to provide a fully integrated network. Base station controllers typically control a small number of base stations and are linked either by a wire land line or often a short-range microwave link.

2.1.2 Handoff procedure

First generation cellular networks follow network-initiated handover. During the call, the network entity base station (BS) measures the signal strength from a mobile device as received at the serving cell and passes this to MSC (Mobile Switching Center). As the signal-to-noise ratio begins to fall below a certain threshold, the nearby cells are requested to perform signal strength measurements. The scanning receiver within the BS measures the signal strength of the MSs in the neighboring cells and reports it to MSC. If the signal received is better in another base station, and if the MSC decides that a handover is necessary, it instructs the neighboring BS to allocate a channel for the mobile. A second voice path is then set up and bridged across to the existing one in preparation for the handoff. Once this is complete, the system or network generates a handoff instruction and sends it over the FOCC (Forward Control Channel). This handoff instruction carries information that includes the new channel to be used along with the SAT (Supervisory Audio Tone). The mobile accepts the information and sends an ST (Signaling Tone) for 50 ms. It then turns off the reverse channel transmission and re-tunes to the new channel. Once it reaches the new channel, the mobile turns its voice transmitter channel transmission on and retransmits the SAT of the new base station. This acts as the confirmation that the new link has been established. Accordingly, when the new base station receives this, it informs the mobile switching center that the handoff has been successfully accomplished. The new base station then informs the old base station to release the channel the mobile was previously using and make it available for further use. Thus, the original voice path via base station A (BS A) gets rerouted via base station B (BS B) as shown in Figure 2.2.
2.2 Cellular 2G mobility

Variety of second generation digital cellular systems have been developed at different parts of the world during early part of 1990. These include GSM/DCS1800/PCS1900 standard in Europe, the PDC standard in Japan, and the IS 54/136 and IS-95 standards in the USA. In this section, I briefly discuss about two second generation mobile systems namely, GSM (Global System for Mobile communication) and IS-95 and highlight their handover procedures.

2.2.1 GSM

GSM (Global System for Mobile communication) system is considered to be a 2G cellular system that is a natural evolution from the TACS (Total Access Communication System) mostly used in European countries.

2.2.1.1 System architecture

Figure 2.3 shows a generalized architecture for GSM. The main elements of the system are the base transceiver station (BTS), the base station controller (BSC), the mobile switching center (MSC) and the registration and authentication components. BTS and BSC form the radio access network (RAN) part of the system. Registration functionalities are provided by home location register (HLR), the visitor location register (VLR). Equipment identity register (EIR) checks the validity of a mobile by checking its international mobile equipment identity (IMEI). Home location register (HLR) contacts the Authentication Center (AuC) in order to authenticate the mobile. These terms are defined in the glossary as part of Appendix B.

A mobile node can be associated with several types of identifiers each with its own function. SIM (Subscriber Identity Module) is a small memory card that provides information about the subscriber. IMEI is a fifteen digit number used to identify the equipment and
is used by the EIR to decide if the mobile is properly authenticated. IMSI is like the home address of the mobile. It enables the operator to link the phone number and the subscriber. TMSI (Temporary Mobile Subscriber Identity) is like a Care-of-address for the mobile that is assigned when the mobile visits a network. TMSI changes whenever the mobile changes its network. In order to provide proper authentication to the mobile, the authentication key is stored on the SIM card that is used to compute the cipher key (Kc). Cipher key is used in the encryption algorithm to prevent unauthorized listening to the mobile message.

There are a variety of control channels that are used to provide required functionality to enable the mobiles and BTS to communicate, set up and manage the calls. These could be split into groups, namely three broadcast channels for initial synchronization, three common control channels for initiating calls and three dedicated control channels to manage calls. BCCH (Broadcast Control Channel), Synchronization Channel (SCH), Paging Channel, Slow Associated Control Channel (SACCH), and Fast Associated COntrol Channel (FACCH) are mostly used for handoff management. SCH provides the BTS identification and allows the mobile to get associated with the BTS. BCCH is continually broadcast on the downlink channel by the BTS and contains the information including base station id and frequency allocations. Paging Channel is used to locate the mobile when there is incoming call. Downlink SACCH provides beacon frequencies of the neighboring cells, and uplink SACCH includes measurement report that gives the strength measurements from signals received from neighboring cell beacon transmissions.

However, I discuss the handoff procedure associated with GSM in the next section.

2.2.1.2 Handoff procedure

Unlike handoff in the 1G cellular systems, handover (handoff) in GSM could be either network initiated mobile assisted handoff or mobile initiated network controlled. Network initiated handoff is triggered by the network based on radio subsystem criteria such as RF (Radio Frequency) signal-to-noise ratio, distance from the base station as well as the
network directed criteria such as current traffic load per cell. In order to support handover that is based on signal-to-noise ratio criteria, MS takes the radio measurements from the neighboring cells and report these to the serving cell on a regular basis. When the network determines that there is a need for handover, appropriate handoff procedures are followed.

Connections in GSM may be handed off between the radio channels in the same cell, between channels in different cells under the same BSS coverage or between the cells under the coverage of different BSSs and even different MSCs. Based on the types of movement, the handoffs can be categorized into two types, namely internal connection handoff and external connection handoff. As part of the internal connection handoff technique, BSS may autonomously handle the connection handoffs in the same cell or between the cells under its own coverage. The MSC is involved in managing the handoffs that need to take place between cells under the coverage of two different BSSs. This process is called external connection handoff. In general, when BSS indicates that an external handover is required, the decision of when and whether an external handover should occur is taken by the MSC.
The MSC uses the signal quality measurement information as reported by the mobile stations that are pre-processed at the BSS for external handover determination. Details of connection handoff are discussed by Rahnema [Rah93].

Figure 2.4: Types of GSM handover

Figure 2.4 shows an example of GSM-based mobility. It illustrates intra BSC, inter BSC and three different types of inter-MSC handover, namely anchor-to-relay, relay-to-relay and relay-to-anchor.

MSC acts like an anchor agent during the duration of the call. During the external handover, the original MSC handling a call keeps control of the call when the mobile is handed over to the target MSC or even to the subsequent MSC. When the BSS performs an internal connection handoff, it informs the MSC when the process is completed. Either the mobile or BSS indicates the impending need for connection handoff by way of FACH (Forward Access CHannel). The BSS usually monitors the quality of radio signal received and transmits such results to the MSC, that keeps a more global view on the radio channels belonging to its BSSs. The MSC may also initiate the need for a connection handoff in an
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

attempt to balance out traffic load in the network.

Table 2.2 shows how different mobility related tasks such as radio channel measurement, handover request and confirm handover are taken care of by the network elements: MS, BS and MSC in a distributed manner.

<table>
<thead>
<tr>
<th>Task</th>
<th>MS</th>
<th>Old Base Station</th>
<th>New Base Station</th>
<th>MSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radio channel measurements</td>
<td>Make periodic measurements on current and neighboring channels Send results to BS</td>
<td>Monitor backwards channels Give measurements order to MS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Start measurements. Send results to BS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issue handover request</td>
<td>Send measurements results to MSC Request Handover Evaluate handover requests. Request handover</td>
<td></td>
<td>Evaluate handover request. Inform new BS</td>
<td>Inform new BS</td>
</tr>
<tr>
<td>Confirm/Disconfirm handover</td>
<td>Accept, block, delay Handover request</td>
<td></td>
<td>Permit, drop, delay handover</td>
<td></td>
</tr>
</tbody>
</table>

2.2.2 IS-95

Interim Standard Number 95 (IS-95) was developed by Qualcomm that got rebranded as CDMAOne in 1997. IS-95 is based on spread-spectrum technology platform that enables multiple users to occupy the same radio channel or frequency spectrum at the same time. IS-95 has two distinct versions, IS-95A and IS-95B. IS-95B provides additional ISDN-like data rates.
2.2.2.1 System architecture

Figure 2.5 shows the system architecture of IS-95 system. It is based on TR-45 [TIAb] reference model. The main components of this architecture are mobile station, base station, mobile switching center, home location register, visited location register, authentication center, and equipment identity register. The base station consists of base transceiver system (BTS) and base station controller (BSC). Besides the switching functions, the MSCs behave as the anchor agents during handoff and help to direct voice traffic. In terms of their functionality, handoff related MSCs can be categorized as anchor MSC, border MSC, candidate MSC, originating MSC, target MSC, serving MSC and tandem MSC. For example, anchor MSC is the first MSC that gives radio contact in a call, a candidate MSC could possibly accept a call or a handoff, serving MSC currently provides service to a call, a tandem MSC provides trunk connection for a call in which a handoff has occurred and a target MSC is the MSC selected for a handoff.

The channels in CDMAOne can be split into forward link channel and reverse link channel. For forward link channel, there are four types of coded channels originating from the base station: pilot, sync, paging, and traffic. For the reverse link channel, there are two types of coded channels originating from the MS: access and traffic. The pilot, sync, paging, and access channels carry the necessary control data while the traffic channels carry digital voice. The pilot channel (PC) provides the MS with a beacon, timing and phase reference and signal strength for power control. The pilot channel is transmitted continuously by each sector of a base station. A mobile uses the pilot channel to identify a cell and can identify the strongest sector within a cell based on the measurement of signal-to-noise ratio of the pilot signal. This channel is useful to initiate the handoff operation. The sync channel (SC) provides the MS with critical time synchronization data. The message on the SC contains information necessary for the MS to align its timing with the pilot channel. The mobile uses the sync channel to discover the network and its parameters. The paging channel contains messages with parameters that the MS needs for access and paging. The messages
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

Convey system parameters, access parameters, and the neighbor list. This channel is used to communicate with the MS when there is no call in progress. Paging channel can be used to locate a mobile.

For the reverse link, there are two basic channels, namely the access channel and reverse traffic channel. Mobile uses access channel to communicate with the base station when no traffic channel has been set up and reverse traffic channel to send data and control signals to the base station.

![IS-95-based architecture](image)

Figure 2.5: IS-95-based architecture

### 2.2.2.2 Handoff procedure

CDMAOne supports three main types of handover: soft handover, softer handover, and hard handover.

**Soft handoff** involves an inter-cell handoff and is a make-before-break connection. The connection between the mobile and the network is maintained by several base stations during the process. A soft handoff occurs only when the old base station and the new base
stations are operating on the same CDMA frequency channel. The mobile communicates simultaneously with these BSs until it is clear that only one BS is required.

**Softer handoff** is an intra-cell handoff occurring between the sectors of a base station and is a make-before-break type. It takes place only at the serving base station.

**Hard handoff** process is functionally break-before-make type. The continuity of radio link is not maintained during the hard handoff. Hard handoff takes place when the mobile is switched between two BSs using different radio channels. Additionally, it is usually meant to enable a mobile to handoff from a CDMA call to an analog call. It is implemented in the areas where a mobile is subjected to handoff from a CDMA network to non-CDMA network because of non-availability of CDMA service in that area.

The handoff process begins when a mobile unit detects a pilot signal that is significantly stronger than any of the forward traffic channels assigned to it and starts to discover a new candidate point of attachment (PoA) where it can connect to. When the mobile detects the stronger pilot signal the following sequence usually takes place: The mobile sends a pilot strength measurement message to the base station and instructs it to initiate the handoff process. The network then sends a handoff direction message to the mobile unit directing it to perform the handoff. On the execution of the handoff direction message, the mobile unit sends a handoff completion message on the new reverse traffic channel.

There are basically two kinds of signaling protocols to take care of mobility related signaling in IS-95: ANSI-634 (American National Standard Institute-634) and ANSI-41. ANSI-634 takes care of signaling between BSC and MSC, and ANSI-41 takes care of the signaling between MSC/VLR and HLR and MSCs.

Figure 2.6 shows the flow for mobile controlled handoff based on IS-41. The handoff process begins with the handoff measurement procedure that determines which of the adjacent systems the mobile should switch to. IS-41-based inter-system handoff consists of two types of handoffs, namely *handoff-forward* and *handoff-back*. Handoff-forward defines a type of inter-system handoff where a mobile is handed off from one mobile switching cen-
ter (MSCA) to another (MSCB), where MSCA is the serving MSC. Handoff-back is the process of handing back to the serving MSC.

IS-95 also offers automatic roaming functionality by which it allows a mobile to originate a call in the visited system and receive the call destined for the roaming subscriber. Three basic processes namely, registration, call origination and call delivery features constitute the roaming function.

### 2.3 Cellular 3G mobility

The main vision of third generation mobile networks (3G networks) is to provide ubiquitous wireless network access that can support voice, multimedia and high-speed data communication. IMT-2000 (International Mobile Telephone-2000) standards were developed by International Telecommunications Union Radio Communication (ITU-R) and International Telecommunication Union Telecommunications sector (ITU-T) to define the operations of 3G networks. One of the main attributes of IMT-2000 is the introduction of wireless wide-
band packet-switched data services for wireless access to Internet with speeds up to 2 Mb/s. IMT-2000 proposed ten different multiple access schemes. Two of these schemes were based on TDMA and the rest eight were based on CDMA. I describe two of these access schemes namely, WCDMA and CDMA2000.

2.3.1 WCDMA

WCDMA (Wideband Code Division Multiple Access) is the natural evolution of GSM-based systems. WCDMA networking specification was created by 3GPP (Third Generation Partnership Project). Compared to 2G version of cellular or CDMA2000, WCDMA uses a much wider spectrum of one 5 MHz channel for both data and voice providing data rate up to 2 Mb/s.

2.3.1.1 System architecture

WCDMA-based 3G networks consist of two major parts: the radio access network (RAN) and the core network (CN). The radio access network consists of base station and radio network controller (RNC). The base station (BS) in 3G (also known as node B) is an interface between the network and WCDMA air interface. The BS is responsible for taking care of channel coding, interleaving, rate adaptation and processing of the air interface. The RNC (Radio Network Controller) acts as an interface between the base station and the core network. It is responsible for controlling the radio resources. The core network in 3G network consists of two domains: a circuit switched (CS) domain and packet switched (PS) domain. Core network consists of components such as VLR, HLR, and MGW on the circuit switched side and SGSN (Serving GPRS Support Node) and GGSN (Gateway GPRS Support Node) on the packet switched side. Detailed description of these functional components can be found in [Tac02].

Figure 2.7 shows the system architecture of a WCDMA network. It shows the components of both packet switched and circuit switched parts of the network.
2.3.1.2 Handoff procedure

In this section, I describe the handoff procedure associated with WCDMA. I briefly discuss the mobility management of WCDMA system, both for circuit switched and packet switched domains. Mobility management for circuit switched domains in WCDMA consists of call origination, termination and handover control. In WCDMA, the mobile node is also called UE. As part of location updating process during attach procedure, a Temporary Mobile Subscriber Identity (TMSI) is assigned to the UE (User Equipment) that is identified by an international mobile subscriber identity (IMSI). IMSI is allocated permanently to each UE (User Equipment). However, every time the mobile changes its network, it obtains a new TMSI and it binds it to an IMSI. This process can be defined as the general binding update process of the handoff procedure.

The handover decisions are generally handled by the RNC (Radio Network Controller). RNC continually monitors information regarding the signals being received by both the mobile node and node B. When a particular link quality falls below a given level and another better radio channel is available, the mobile initiates a handover. As part of this
monitoring process, the UE measures the received signal code power (RSCP) and received signal strength indicator (RSSI) and information is then returned to the node B and hence to the RNC on the uplink channel.

There are several cases of handovers in WCDMA based on the movement pattern of the UE with respect to the cell and other nodes (e.g., Node Bs, RNCs) within the UTRAN (UMTS Terrestrial Radio Access Network). WCDMA-based handoff can be categorized into three types: soft handover, softer handover and hard handover.

**Soft handover:** Soft handover takes place when the operating frequency remains same between the neighboring cells. Several kinds of soft handover types are discussed here.

1. **Intra Node B/intra RNS (Radio Network Subsystem):** This handover type is done if the UE moves from one cell to another cell where both belong to the same Node B.

2. **Inter Node B/inter RNS/intra SGSN:** If the UE moves from a cell in one Node B to a cell in another Node B that belong to different RNS.

3. **Inter Node B/inter RNS/inter SGSN:** In this case, the UE moves from a cell in one Node B to a cell in another Node B that belongs to a different RNS. The Node Bs are connected to different RNCs and those two RNCs are also connected to different SGSNs. In this case, the UE may even move between two SGSNs.

4. **Inter GGSN:** GGSN is the layer 3 point of attachment that acts as the default router. Unless the mobile moves between GGSNs the layer 3 identifier of the mobile does not change.

**Softer handoff:** When the UE moves between two different sectors on the same base station it is called softer handoff. Since the processing elements are shared, it enables a softer handoff to be accomplished more easily than a soft handoff.

**Hard handover:** Hard handover (or inter-frequency handover) is only needed if the mobile needs to change its frequency after the handover or interface does not exist between
two RNCs in case of a soft handover (inter Node B/inter RNS). A frequency change may take place as a result of change of W-CDMA cell level i.e., from a macro cell to a satellite or another change of the radio access technology (inter RAT handover), for example from UMTS to a WLAN or GSM network. A hard handover occurs quite rarely and differs a lot from the soft handover types described above.

### 2.3.2 CDMA2000

CDMA2000 builds on CDMAOne to provide an evolution path to 3G. Like WCDMA, it supports both kinds of access: FDD-based (Frequency Division Duplex) DS/CDMA and TDD-based (Time Division duplex) T/CDMA. There are several versions of CDMA systems available, namely CDMA2000 1X, CDMA 1xEV-DO (Evolution Data Only), and CDMA 1xEV-DV (Evolution Data and Voice).

#### 2.3.2.1 Architecture

Compared to WCDMA, it uses a spectrum of 1.25 MHz per channel. CDMA2000 1X doubles the voice capacity of CDMAOne system and also supports high-speed data services. It can support peak data rate of up to 153 kb/s. CDMA 1xEV-DO is an evolution of CDMA 2000 that is designed for data-only use and it provides a peak rate capability of over 2.45 Mb/s on the downlink. On the other hand, CDMA2000 1xEV-DV is an evolution of CDMA 2000 that can simultaneously transmit voice and data. The peak data rate is 3.1 Mb/s on the forward link and reverse link is limited to 384 kb/s.

In order to be able to support packet based services more efficiently, it upgrades the existing network elements of IS95, namely BTS and BSC and adds many new network elements, like PDSN (Packet Data Serving Node), AAA, and home agent (HA). BSC is equipped with additional IP routing functionality and new PDSN establishes, maintains, and terminates point-to-point sessions with the subscriber and initiates authentication, authorization and accounting (AAA) for the mobile station client to the AAA server. In
addition, it also augments the functionality of HLR and VLR. For an architecture based on CDMA2000 in conjunction with legacy network components such as MSCs, HLRs, signaling between BSC and MSC are taken care of by TIA 2001 [TIAa] standards, and signaling between MSC, VLR and HLR and between MSCs are taken care of by ANSI-41 standards. TIA-2001 describes the overall system functions including services and features required for interfacing a Base Station with the Mobile Switching Center (MSC), with other Base Stations (BSs), with the Packet Control Function (PCF) and for interfacing PCF with the Packet Data Service Node (PDSN). MN usually initiates a data call via BTS which is the first point of network attachment. The BSC responsible for this BTS (Base Transceiver Station) forwards the call to the associated PCF (Packet Control Function). The PCF selects a PDSN based on certain unique characteristics of the mobile and establishes a GRE (Generic Routing Extension) tunnel with the PDSN. At that point MN initiates a PPP session that gets terminated at the PDSN. Thus, there is a single hop IP connectivity between the mobile and the PDSN (Packet Data Serving Node).
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

2.3.2.2 Handoff

Compared to IS-95, CDMA2000 introduces additional levels of thresholds to limit the frequent handovers. By limiting the unnecessary handovers, it increases the system capacity for the data traffic. Like WCDMA, CDMA2000 supports three types of handoff operations: hard handoff, soft handoff, and softer handoff. Figure 2.8 shows a hierarchical network architecture with all the networking elements of a CDMA2000 network. A node’s mobility is taken care of at each level of the hierarchy, namely at BTS level, PCF level, and PDSN level. There are several movement scenarios that can take place within this hierarchy. When the mobile moves between two BTSs, as long as the movement is confined to the PDSN the mobile does not need to set up a new PPP session. But if the mobile moves between two BTSs that are controlled by two different PCFs, each PCF may choose a new PDSN in the hierarchy. This will require to terminate the PPP session and start a new one. When the mobile chooses a new PDSN to re-establish a new PPP session, it obtains a new layer 3 identifier such as IP address. During packet-based communication, a mobile node is identified with a simple IP addressing scheme or mobile IP. In case of simple IP addressing scheme, the mobile obtains the IP address from a DHCP server that usually co-locates with the PDSN. This identifier changes as the mobile moves between the PDSNs. However, if the mobile uses mobile IP-based approach, this IP address does not change. In case of IP address change, layer 3 mobility can be taken care of by mobility protocols like Mobile IP at PDSN layer as discussed in Section 2.4.1.

2.4 4G Networks

4G refers to the fourth generation of cellular wireless and is an evolution of 3G networks. 4G networks provide additional features such as higher data rate up to 100 Mb/s, seamless mobility support across heterogeneous access networks, secured IP-based communication, QoS support for many multi-media services such as mobile TV, MMS (Multimedia
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

Messaging Service), DVB (Digital Video Broadcasting). 4G networks are based on several access technologies, namely OFDMA (Orthogonal FDMA), Flash-OFDM, SC-FDMA (Single Carrier-FDMA), MC-CDMA (Multi-carrier CDMA). LTE (Long Term Evolution) does support OFDMA, WiMAX (IEEE 802.16) supports Flash-OFDM, and IEEE 802.20 supports MC-CDMA as the access technique. However, I describe one specific 4G candidate network, LTE, and its mobility features in this section.

2.4.1 Evolved Packet System

2.4.1.1 System Architecture

EPS (Evolved Packet System) represents the very latest evolution of UMTS standard. EPS has been defined as part of release 8 within the 3GPP standards bodies. It consists of two parts: LTE (Long Term Evolution) and SAE (System Architecture Evolution). LTE takes care of the evolution of radio interfaces and SAE focuses on core network architecture evolution. Figure 2.9 shows the evolved core architecture (EPC). I briefly describe some of the functional components of LTE and EPS.

LTE, which is also known as Enhanced UTRAN (E-UTRAN) defines a packet optimized access network that can efficiently support IP-based real-time and non-real time services providing performances comparable to circuit switched networks. Unlike UTRAN, E-UTRAN relies on a fully shared radio resource allocation scheme that allows maximizing resource usage by combining all radio bearers on a shared high bit rate radio channel.

The goals of E-UTRAN are to provide reduced latency, higher data rates, improved system capacity and coverage as well as reduced cost for the operator and packet optimized radio access technology. The most important component of E-UTRAN is eNode-B. Unlike UMTS, there is no separate RNC, but RNC functionality has been integrated into eNode-B.

SAE also defines a simplified core network composed of one packet domain that can support all packet switched services and inter-working capabilities with traditional PSTN. EPS represents a migration from the traditional hierarchical system architecture to a flat-
tended architecture that minimizes the number of communication hops and distributes the processing load across the network. SAE is also known as EPC (Evolved Packet Core). EPC is composed of several functional entities, namely MME (Mobility Management Entity), serving gateway (S-GW), PDN (Packet Data Network) gateway, and PCRF (Policy and Charging Rules Function). MME is in charge of all the control plane functions and provides features such as security procedures, terminal-to-network session handing, idle terminal location management. The serving gateway (S-GW) is the termination point of the packet data interface towards E-UTRAN and serves as a local mobility anchor supporting intra E-UTRAN mobility and mobility with other 3GPP technologies such as 2G-based GSM and 3G-based UMTS. PDN GW (Packet Data Network Gateway) is the termination point of the packet data interface towards the packet data network. PDN GW supports policy enforcement features, packet filtering and enhanced charging support. PDN-GW also acts like a home agent.

EPS provides much better performance compared to the UMTS of the previous releases.
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

A peak data rate of up to 100 Mb/s downlink and 10 Mb/s up-link can be obtained with EUTRAN. It offers a radio access network latency of 5 ms in user plane and 100 ms for control plane. Latencies for inter RAT (Radio Access Technology) are restricted to 300 ms for real-time traffic and 500 ms for non-real-time traffic.

2.4.1.2 Handoff procedure

LTE supports mobility both at radio layer and network layer. It supports UE-assisted network controlled hard handover when the UE is in active mode. There are two types of handoff defined for LTE: Intra-MME and Inter-MME. During Intra-MME handoff, the mobile moves between two eNBs that are connected to the same MME. Inter-MME defines the handover when the eNBs are controlled by two different MMEs. Two eNBs could be served by two different serving gateways (S-GWs).

LTE supports mobility with both 3GPP and non-3GPP access systems. Mobility across 3GPP access systems is called local mobility and mobility across non-3GPP access systems is called global mobility. Non-3GPP access systems include wireless systems, such as IEEE 802.11 and IEEE 802.16. Both GTP-based (GPRS Tunneling Protocol) mobility and mobile-IP-based mobility are used to support local mobility. However, only mobile IP-based mobility is used to support global mobility.

I describe below the specifics of a handoff process involving E-UTRAN. Handoff management within LTE network involves several stages, namely discovery, measurement, handover preparation and finally, the handover execution. The user equipment (UE) that is equivalent to a MH (Mobile Host) can be identified with a list of identifiers that are used to assist handoff. IMSI and IMEI provide subscriber ID and equipment ID, respectively. M-TMSI (M-Temporary Mobile Subscriber Identity) is used to identify the UE within the MME. GUTI (Globally unique temporary UE identity) is allocated by MME. GUTI identifies a globally unique MME and the UE within the MME. The UE is assigned S-TMSI (S-Temporary Mobile Subscriber Identity) which is a shortened version of the GUTI and
is used to locate the UE. UE is also assigned an ID that uniquely identifies the UE in a tracking area. An eNodeB sends the tracking area identifier (TAI) on Broadcast Control Channel (BCH). Upon power on the UE performs a cell search, discovers the EPS/LTE access system and performs access system and network selection. After the cell is discovered, it attaches to the radio network and is authenticated with the MME. After the successful authentication, the MME registers with the HSS (Home Subscriber Server) as the serving UE in the HSS. Network layer configuration is done by the PDN-GW so that mobile can configure itself with an IP address and the default router.

Figure 2.10 describes intra-MME handover process where the UE is handed off from eNB1 to eNB2 that are part of the same MME. It shows the signaling between eNBs, MME and signaling gateway (S-GW).

![Diagram](image)

**Figure 2.10: Intra-MME handoff call flow**

Based on certain policy, the UE sends a measurement report to the source eNB or serving eNB. The source eNB makes a decision based on the measurement report and radio resource management (RRM) function to handover the UE to the target eNB. The source
eNB sends a handover request to target eNB and passes the necessary QoS information it presently supports for the UE. This ensures that the target eNB configures the required resources for the UE when it arrives in the target network. A GTP-U tunnel is also set up between the source eNB and target eNB during this time. The mobile is then assigned the radio resources by the target eNB. During this time, the data destined to UE gets buffered in the source eNB. Once all the details of the UE such as UL (Uplink) PDCP sequence number receiver status and DL (Down Link) sequence number transmitter status are known to the target eNB, it receives the buffered data from the source eNB and buffers it locally for further delivery to the UE at a later point. Once the UE connects to the target eNB successfully, the target eNB informs the MME to do the PATH SWITCH. On receiving the PATH switch request, the MME communicates with the S-GW to modify the bearer through modify bearer request. S-GW transmits the end of marker packets to the source eNB before starting transmission on the new path. This helps the source eNB to stop forwarding packets to the target eNB. Target eNB delivers the buffered packets before delivering packets on the new DL path. After the new path has been established, source eNB releases its resources. Thus, the new data is delivered to the UE via target eNB. Buffering at the source eNB and target eNB help to reduce the packet loss during handoff.

2.5 IP-based mobility

IP-based mobility management techniques can be implemented at several layers of the protocol stack, such as network layer, transport layer, and application layer. IP-based mobility protocols can be used to take care of mobility for 3G- and 4G-based systems. MIPv4 (Mobile IPv4) [Per02c] and its several variants, namely MIP-RO (MIP with Route Optimization), and MIP-RR (MIP with Regional Registration) [Per02b], MIP-LR (MIP with Location Register) [JRY+99], MIPv6 [JPA04], MOBIKE [SE06] are a few of the network layer mobility protocols that were defined within the IETF. Cellular IP [CGK+00],
HAWAII (Handoff Aware Wireless Access Internet Infrastructure) [RLS+00], Proxy MIPv6 [GLD+08], IDMP (Intra Domain Mobility Protocol) [DDM+02] are the network layer micro-mobility protocols suitable for intra-domain mobility. Intra-domain mobility refers to a movement scenario when the mobile’s movement is confined to the administrative domain. MSOCKS [MB98], TCP-Migrate [SB00], and SCTP (Stream Control Transport Protocol) [K+03] have been designed to take care of mobility at the transport layer. SIP-based mobility [SW00] takes care of mobility by means of application layer signaling, such as SIP (Session Initiation Protocol) [RSC+02]. HIP (Host Identity Protocol) [MN06] defines a shim layer between the network layer and transport layer to provide terminal mobility in a way that is transparent to both the network layer and transport layer.

I have provided a survey of the related mobility protocols and the issues in [DACS02], [DVC+01]. I have also experimented with several mobility protocols, namely, MIPv4, MIPv6, SIP-based mobility, MIP-LR, and ProxyMIPv6 and verified that these mobility protocols in their current form are not adequate to meet the delay, and packet loss performance requirements of real-time traffic [DKZ+05], [DMD+07] and hence, these protocols will benefit from overall systems optimization.

I briefly describe some of these IP-based mobility protocols and have categorized them into network layer macro mobility, network layer micro mobility, application layer mobility and transport layer mobility.

### 2.5.1 Network layer macro mobility

Network layer mobility can be categorized into two types: *macro* and *micro*. Macro mobility mechanism takes care of global mobility where the mobile moves between administrative domains. I describe two types of network layer macro mobility, namely Mobile IPv4 and Mobile IPv6.
2.5.1.1 Mobile IPv4

Mobile IP is a mechanism developed for the network layer to support mobility [Per02a]. Originally it was intended for travelers with laptops to provide portability, and later has been adopted by the wireless community. It supports transparency above the IP layer, including the maintenance of active TCP connections and UDP port bindings. A mobile host is identified by a node identifier such as fixed IP (home IP address). When the mobile host connects to a visited network that is different than the one where its IP address belongs to, its home network forwards packets to the mobile. A router (or an arbitrary node) that is usually known as the home agent on the user's home network forwards the packets. There are two different methods to deliver the packets to a mobile host when it is on a foreign network. With the first method, the mobile host adopts a second (temporary) IP address known as the care-of address (CoA) and registers it with its home agent. When the home agent receives a packet for this user, it encapsulates the packet in another IP packet with the care-of address as the destination address and sends it to the foreign network [Per96a], [Per96b]. Encapsulating a packet within another packet until it reaches the care-of address is known as tunneling. Note that encapsulation adds between 8 and 20 bytes of overhead, which can be significant for voice packets of this size.

The care-of address in the first method is said to be co-located and it can be acquired via services like DHCP (Dynamic Host Configuration Protocol) [Dro97] or its optimized version such as DHCP with rapid commit [P KB05] in a local area network or via PPP [Sim94] in a point-to-point networking environment. With the second method, the mobile host first registers with a foreign agent (FA) in the network it is visiting. The foreign agent sends (registers) its address to the mobile host’s home agent as the care-of address of the mobile host. Packets that are intended for the mobile host are sent to the foreign agent after the home agent encapsulates them with the IP address of the foreign agent. After decapsulating these packets, the foreign agent delivers them to the mobile host.

Figure 2.11 shows the functional details of Mobile IPv4. In this specific figure the
mobile node moves from subnet 1 to subnet 2 and in the process changes its layer 2 point of attachment, layer 3 point of attachment and either reconfigures itself with a new care-of-address from a DHCP server or uses FA’s address as its new CoA.

![Diagram of Mobile IPv4](image)

**Figure 2.11: Mobile IPv4**

For the methods outlined above to be able to work, a mobile host needs to be able to learn that it has moved from its home network to a foreign network. For this purpose, home agents and foreign agents advertise their presence periodically in their own broadcast domains. A mobile host can also solicit agent advertisements in case these advertisements are absent. Packets from the correspondent host must first travel to the home agent that are later forwarded to the mobile host either by way of the foreign agent or directly. Packets from the mobile host do not have to traverse the home agent; the mobile host sends them as usual with its home IP address as the source address, which is known as triangular routing.

Routing of all incoming packets via the home network may cause additional delays and waste of bandwidth. However, if the correspondent host knows where the mobile host is, it can send packets directly to the care-of address of the mobile host. This is achieved...
by route optimization [Per02a] process that enables the mobile to send mobility binding updates directly to the corresponding host. Binding updates are sent from a home agent upon request of the mobile, or can be sent upon receiving a warning message from a foreign agent if the mobile host changes location during a communication session. In the second case, the former foreign agent will keep forwarding packets to the new foreign agent until the correspondent host updates its mobility binding cache (known as smooth handoff).

Another optimization proposed is regional registration [CMP03], where the mobile locally registers within a visited domain. In base Mobile IP, a mobile host is required to register with its home agent each time it changes care-of address, thus causing signaling delay for the registration if the mobile host is far away from its home agent. Regional registration attempts to decrease the number of home registrations by maintaining a hierarchical structure of the foreign agents. As long as a mobile host’s foreign agent is located hierarchically under a so-called gateway foreign agent (GFA), it is unnecessary to relay registration messages back to the home agent since the home agent has already registered the GFA’s address as the care-of address. To make Mobile IP handoffs (i.e., the registration process) more suitable for real-time and delay-sensitive applications, Malki et al. [Mal04] proposed two additional methods. With the first of these methods, called the Network Assisted, Mobile and Network Controlled (NAMONC) handoff method, the mobile host is informed (assisted) by the network that a layer 2 handoff is anticipated. It proposes to use simultaneous bindings (multiple registrations at a time) in order to send multiple copies of the traffic to potential points of attachment before the actual movement. The other method, called the Network Initiated, Mobile Terminated (NIMOT) handoff method proposes extensions to the base mobile IP so that foreign agents can utilize information from layer 2. Specifically, foreign agents use layer 2 triggers to initiate a pre-registration prior to receiving a formal registration request from the mobile host. Both methods assume considerable involvement of information from layer 2.

I have proposed many of the Mobile IP related optimization techniques in Chapter 5.
2.5.1.2 MIP-LR

Mobile IP with Location Registers (MIP-LR) avoid encapsulation of packets [JRY+99] and provide survivable features in case of failure of location registers. In MIP-LR, each subnet may contain a host that functions as a visitor location register (VLR) and/or a host that functions as a home location register (HLR). Each mobile host can be served by multiple HLRs. Each VLR advertises its presence on its local subnet using agent advertisement messages similar to Mobile IP. When a mobile host is located at its local subnet, it is not registered at either the HLR or the VLR. When the mobile moves to a foreign network it obtains a care-of address (CoA) from the pool of addresses that VLR has. The mobile host registers with the foreign VLR using the CoA it has obtained, which in turn relays the registration to the mobile host’s HLR. The HLR returns a registration reply containing the allowed lifetime for this registration; the VLR records the mobile host’s CoA and the lifetime and forwards the reply to the mobile host. A correspondent host, wishing to send a packet to the mobile host for the first time, issues a query to the HLR, which returns the mobile host’s CoA as well as the remaining registration lifetime. The correspondent host then directly sends the packet to the mobile host’s CoA. The correspondent host caches a binding for the mobile host’s CoA and uses this binding for subsequent packets destined to the mobile host. The correspondent host must refresh its binding cache by querying the HLR again before the mobile host’s remaining registration lifetime expires. In MIP-LR, unlike Mobile IP, HLR can be geographically distributed anywhere. As part of my thesis, I have implemented an extension of MIP-LR using application layer module that does not need any kernel changes. Having an application layer module, it allows the mobile to use a policy-based approach to trigger MIP-LR for certain types of application. Figure 2.12 shows the functionalities of Mobile IP with location registers when the mobile moves from one subnet to another and in the process changes its layer 2 point of attachment and layer 3 point of attachment. In this case, there is no foreign agent in the visited network and it is also not a requirement that the location register needs to be in the home network.
After a mobile host moves, if the mobile host was previously registered at some other foreign VLR, the new VLR de-registers the mobile host at the old VLR. The de-registration is required so that the mobile host’s old CoA can eventually be released for use by some other mobile host. If a VLR runs out of CoAs temporarily, it can still issue its own IP address as a CoA, and, when a mobile host registers using this CoA, inform the HLR accordingly.

### 2.5.1.3 Mobile IPv6

IPv6’s [DH98] increased address space and inherent support for security and auto-configuration have made it an attractive candidate to support mobility for the next generation Internet. Mobile IPv6 is the protocol to support mobility for IPv6 nodes. Since address auto-configuration is standard part of MIPv6, MH will always obtain a CoA routable to the foreign network. Thus, there is no need to have a foreign agent (FA) in MIPv6. When the mobile node moves to a new foreign network it acquires a temporary care-of-address using
stateless auto-configuration [TN98] or via DHCPv6 [DEB+03].

Figure 2.13 shows the functional components of Mobile IPv6. Unlike Mobile IPv4, the visited networks do not have any foreign agent. MIPv6's route optimization feature also enables direct data delivery from the CH to the mobile node.

![Figure 2.13: Mobile IPv6](image)

Although Mobile IPv6 is defined as a network layer approach and one needs to install MIPv6 stack so as to support mobility in IPv6 space, in future, any standard operating system will come with inherent mobile IPv6 support.

While mobile IPv6 provides a way of making sure the uniqueness of an address as the mobile moves to a new router space it also adds delay for the binding update and binding acknowledgement like mobile IPv4. However, compared to regular Mobile IP, there are inherent advantages in MIPv6. Route optimization is a standard feature of MIPv6, thus, there is no need for the CH to be equipped with additional software like MIP-RO (Mobile IP with route optimization). MH sends binding update directly to the correspondent host (CH) and makes use of the home address destination option as part of the binding update.
This allows the correspondent host to keep a binding cache that maps the care-of-address of the mobile with the mobile’s home address. For the on-going traffic it avoids triangular routing, thus packets from CH to MH need not be encapsulated but are sent directly to MH with its COA as the source route. However, when a new CH needs to communicate with the mobile for the first time, the packets from CH need to travel to HA and get tunnelled to the mobile host. As the mobile moves during the packet transfer process, the subsequent packets are tunnelled directly to the mobile host without being routed via the home agent.

2.5.2 Network layer micro mobility

There are a few network layer micro mobility protocols that are meant to optimize mobility when mobile’s movement is confined within a domain. These protocols avoid the overhead associated with tunneling over the air, and reduce the signaling overhead when the mobile’s movement is confined to a domain. I describe a few of those network layer micro mobility protocols, namely, cellular IP, HAWAII, IDMP and ProxyMIPv6. I mainly focus on the general mechanisms that they use to optimize the mobility.

2.5.2.1 Cellular IP

Cellular IP [CGK*00] is a micro-mobility management protocol. It separates local and wide area mobility, adopts a domain-based approach, and uses Mobile IP for inter-domain (wide area) mobility. Cellular IP isolates the wireless access network from the core of the Internet via a gateway that acts like a foreign agent and deploy network elements (base stations) specialized for mobility management. Isolating the wireless access from the core is necessary since Cellular IP itself provides an IP forwarding engine replacing IP in the wireless access network. This approach reduces the signaling updates and localizes it within a domain.

Figure 2.14 shows how the packets destined to the mobile host are routed through the cellular IP nodes as the mobile moves between cells within the same domain and between
the domains. Packets from a correspondent host are first sent to the mobile host’s home agent and then tunneled to the gateway where they are decapsulated. Hosts are identified by their home addresses inside the Cellular IP cloud. Packets generated by mobile hosts are sent to the gateway and later to the correspondent host. Each cellular IP domain is equipped with a gateway router that periodically broadcasts beacon. This beacon is broadcast to all the cellular IP nodes within the domain and thus each node between the mobile and gateway learns of its neighbor’s address in the uplink and uses that information to route any data traffic destined to the gateway. Cellular IP base stations snoop actual data packets sent from mobile hosts to the gateway to cache the path taken by them (actually, base stations record only the host IP number and the neighboring base station from which the packet was received). To route packets from the gateway to the mobile host, base stations use the reverse of this path. Hosts that have not transmitted packets for a while are removed from the routing cache of the base stations. Location-tracking method of mobile hosts depends on whether the hosts are active or idle. An idle host is one that has not received or trans-
mitted a packet for a specific time. It is adequate to maintain the position of idle hosts as a
distributed paging cache. To achieve this, a technique known as passive connectivity in cel-
lular telephone systems is mimicked in Cellular IP layer. Base stations are geographically
boxed into paging areas, where a paging area may include more than one base station.
Idle hosts send infrequent paging-update packets to the gateway. For active hosts, a dis-
tributed routing cache maintains the exact location of each host. Once a mobile host moves
to another base station during a call, it sends a route-update packet back to the gateway.
New base station(s) record this path accordingly.

2.5.2.2 HAWAII

Handoff-Aware Wireless Access Internet Infrastructure (HAWAII) is another effort to com-
pensate for Mobile IP’s inefficiency in supporting intra-domain mobility [RLS+00]. In that
sense it is similar to Cellular IP’s objective, but unlike cellular IP, HAWAII defines sepa-
rate control messages to set up host routes on the intermediate routers to route the packets
between the domain root router and the mobile host. HAWAII operates on the basic as-
sumption that most user mobility occurs within a domain, therefore optimizing routing
and forwarding for efficient support of intra-domain mobility will complement Mobile IP’s
inter-domain mobility support. Another assumption is that base stations are capable of IP
routing. HAWAII segregates the network into a hierarchy of domains. On the top of hier-
archy for each domain, there is a domain root router. Packets addressed to a mobile host
in a specific domain, first reaches the domain root router and are then sent to the mobile
host. As long as the mobile host moves within a domain, it retains its IP address. Once
the mobile host moves into another domain it is assigned a co-located care-of address, and
the home agent in the home domain tunnels packets to this address. Figure 2.15 shows a
sample architecture for HAWAII.

The path (route) between the mobile host and the domain root router is specific to that
host. It is established during power-up and updated during movement for that mobile host
in the domain root router and pertinent intermediate routers. This information is refreshed periodically by the mobile host, which allows the routers to maintain path state. This idea is similar to Cellular IP and the regional registration in Mobile IP: a physically distant home agent involvement is not desirable each time the mobile host moves. Four path setup methods that can be used to re-establish path states when the mobile host moves within a domain are proposed in this scheme. Two of them forward packets from the old base station to the new base station for a short period (i.e., until the relevant routers update their entries for the specific host). The other two methods do not forward packets during a handoff. Rather, they either bi-cast the packets to two base stations or unicast them for hosts that can simultaneously listen to two base stations. Obviously, HAWAII requires all routers in a domain to be augmented with mobility support so that they are able to handle host-specific path setup messages.
2.5.2.3 TeleMIP

TeleMIP (Telecommunications-Enhanced Mobile IP) is an intra-domain mobility framework which uses two layers of scoping within a domain and is based on IDMP [DDM*02] (Intra Domain Mobility management Protocol). Figure 2.16 shows the basic architecture for TeleMIP. By specifying an intra-domain termination point called mobility agent (MA) it helps to reduce the signaling updates during the movement within a domain. Mobility agent has the similar functionality like a foreign agent but is placed hierarchically at a higher level. This reduces the signaling traffic due to frequent hand-offs within a domain. Mobility agent acts like a domain-wide anchor point similar to gateway foreign agent (GFA) in MIP-RR. Unlike other proposed intra-domain mobility management schemes, IDMP uses two dynamically auto-configured care-of addresses (CoAs) for routing the packets destined to mobile nodes. The global care-of address (GCoA) identifies the mobile node’s attachment to the current domain, while the local care-of address (LCoA) changes with subnet change and identifies the mobile’s attachment in a specific subnet.

Intra-domain updates are sent only up to the MA, which provides a globally valid care-of address (CoA) to the mobile host. TeleMIP reduces the frequency of global update messages since the MA is located at a higher hierarchy than that of subnets, global updates (to home agent, correspondent hosts, etc.) only occur for inter-domain mobility. It also reduces the requirement of additional public addresses (IPv4) by adopting a two-level addressing scheme.

In TeleMIP the network is divided into domains like Cellular IP and HAWAII. Each domain is identified by the mobility agent’s IP address. Domain identifiers are broadcast as part of agent advertisement in each subnet. Each mobility agent is identified with a unique domain identifier. MA’s address can also serve as a domain identifier. Network retains control of MA assignment. When a mobile host first moves into a domain, it obtains a global care-of address (mobility agent’s address) as well as a local care-of address. MA’s global CoA is sent in the registration message to HA. Mobile host also registers itself
(using its local CoA) with its MA. There are foreign agents and DHCP servers organized hierarchically at the subnet level under an MA. They provide the mobile host with a locally-scoped address (LCoA), which identifies mobile location within the domain. The mobility agent provides mobile host with a global care-of address that stays constant as long as the mobile stays within the domain.

Each domain is equipped with at least one mobility agent. However, multiple MAs can be provisioned for load-balancing and redundancy within the domain. Mobile host's location is known only up to the MA-level granularity. A mobile host retains the same MA (global care-of address) within the same domain. All packets from the global Internet are tunneled to the MA, which acts as a single point of enforcement. The MA forwards packets to mobile host, using regular IP routing, by using the local CoA (co-located or FA) as the destination. On subsequent movement within the domain, mobile host only obtains a new local CoA. At that point there is no need to update the home agent or correspondent hosts. However, mobile host updates its MA with its new local CoA. With TeleMIP, if packets
come from outside the domain, they go through the process of encapsulation twice once at the home agent and once at the mobility agent, thus adding delay to the packet delivery process due to additional processing at the mobility agent.

As part of my thesis, I have designed and implemented fast-handoff techniques associated with TeleMIP. I will describe these fast-handoff mechanisms in Chapter 5.

### 2.5.2.4 Proxy MIPv6

Advantage of local mobility management is to optimize many of the functions related to mobility and reduce the number of signaling messages over the air. A candidate mobility protocol called Proxy MIPv6 [GLD+08] has been standardized within the IETF to optimize local mobility management operations. This protocol is designed to take care of local mobility and is controlled by the network elements thereby reducing the load on the mobile nodes and number of signaling messages over the air. PMIPv6 does not use any mobility stack on the mobile node, but rather depends upon the proxies on the edge routers to perform the required mobility functions. These proxies are called proxy mobility agents (PMA) and can co-locate with the edge routers that are often called as media access gateway (MAG). As long as the mobile node moves within the same domain that has the same local mobility agent (LMA), the mobile node assumes that it is in a home link. The PMA is responsible for sending the proper mobile prefix as part of the router advertisement for stateless auto-configuration, or it can also act as a DHCP relay agent for stateful auto-configuration of the mobile nodes.

I briefly describe the operations of Proxy MIPv6. When the mobile node moves from one MAG (Media Access Gateway) to another MAG, and its movement is limited within one LMA, the following mobility related operations are performed: layer 2 movement, detection of new link by way of router solicitation, access authentication, profile verification, proxy binding update and address re-configuration.

Figure 2.17 illustrates the network elements associated with PMIPv6 operation. After
the mobile node connects to the new point-of-attachment as part of the initial bootstrapping process or after the movement to a new domain, access is authenticated with the designated AAA server. During this process, PMA sends the binding update to the LMA (Local Mobility Agent) with the address of the PMA that is specific to the home prefix of the mobile node. In the absence of a pre-existing tunnel, this process helps to set up a tunnel between the LMA and the respective PMA. The mobile node configures its address using the prefix included in the router advertisement and interface-id, which can be assigned by PMA or created by itself. After the movement to the new access network, if the same prefix is advertised by the new PMA, then the IP address of the new mobile does not change. A tunnel is not desirable on the mobile node because it adds extra processing and bandwidth constraints to the wireless hop. PMIPv6-based mobility protocol is preferred when mobility is confined within a domain and wireless service providers do not want to overload the mobile node’s stack by setting up a tunnel between the mobile and the HA.

Figure 2.17: Proxy Mobile IPv6
2.5.3 Transport layer mobility

There are a few transport layer mobility solutions. TCP-Migrate [SB00] proposes a new set of migration options for TCP which provide a pure end-system alternative to network layer solutions. With this extension, established TCP connections can be suspended by a TCP peer and reactivated from another IP address without a third party except for the involvement of dynamic DNS updates for locating the mobile host. However, this approach requires modifying the transport protocol at the end-terminals. MSOCKS [MB98] is another transport layer solution that introduces a proxy in the middle of a network and is built on top of the SOCKS protocol [LGL+96] often used for firewall traversal. Upon movement of the mobile and its address change, the intermediary proxy helps splice the TCP connection. The transport protocol SCTP (Stream Control Transport Protocol) [SXMS00] also adds support for mobility. It has a built-in ADDIP feature [K+03] that provides continuity support when the mobile’s IP address changes. Figure 2.18 shows the splicing of TCP connection when there is a break in the communication due to mobile’s movement.

Figure 2.18: MSOCKs-based mobility
2.5.4 Application layer mobility

Application layer mobility uses the Session Initiation Protocol (SIP) [RSC*02] as the underlying signaling mechanism to provide host-based mobility solution. This mechanism does not depend on the home agent or foreign agent in the network nor does it require to have additional mobility software in the end hosts. Figure 2.19 shows the functional components of SIP-based mobility. Just like other mobility protocols it illustrates a similar scenario where the mobile moves between two different subnets and in the process changes its layer 2 point-of-attachment and layer 3 point of attachment. Just like direct binding updates in case of MIPv6 and MIP-LR, SIP-based mobility management uses Re-Invite signal to the correspondent host as a way of direct binding update that helps to update the binding cache in the correspondent host. Thus, the data is delivered using an optimized direct path. Application layer mobility based on SIP can support variety of mobility mechanisms such as personal mobility, service mobility and terminal mobility. As part of my research, I have experimented with SIP-based application layer mobility protocol for both RTP- and TCP-based applications and have designed several optimization techniques for SIP-based application layer mobility protocol. These optimization techniques for SIP-based mobility management are described in Chapter 5.

2.5.5 Host Identity Protocol

HIP (Host Identity Protocol) [MN06] defines a new protocol layer between inter-networking layer and transport layer to provide terminal mobility in a way that is transparent to both network layer and transport layer. The proposed host identity namespace takes care of some of the important gaps between the existing IP and DNS namespaces. Host Identity namespace consists of host identifiers (HIs). A host identifier is cryptographic in its nature and it is the public key of an asymmetric key pair. Each host identity uniquely identifies a single host. By doing so, IP addresses can be decoupled from the higher layer applications in a secure manner and the end hosts can be authenticated by a public key as they move around. Public
key is one component of an asymmetric cryptographic key pair used as a publicly known identifier for cryptographic identity authentication. By decoupling network and transport layers, applications and transport connections can be made independent of underlying IP address changes thereby enabling alternative solutions to host mobility, host multi-homing, and network address translation. Potential benefit of HIP is that it can be directly integrated with IP security protocols IPSec. Figure 2.20 shows how the end-point address and the locator are separated whenever IP address changes.

Since IPSec security associations (SAs) are bound to host identifiers, not addresses, IP address change does not break the existing transport connections, nor does it trigger a re-establishment of IPSec SAs. HIP-enabled mobility provides optimization technique similar to route optimization technique of Mobile IPv6 by sending direct update to the correspondent host. However, unlike MIPv6, HIP does not have any home network. HIP uses a Readdress packet that is similar to Binding Update for MIPv6. However, unlike MIPv6, it inherently secures the readdressing process. In the HIP architecture, the end-
point names and locators are separated from each other. IP addresses continue to act as locators. The Host Identifiers take the role of end-point identifiers. It is important to understand that the end-point names based on Host Identities are slightly different from interface names; a Host Identity can be simultaneously reachable through several interfaces.

2.5.6 MOBIKE

MOBIKE (IKEv2 Mobility and Multihoming) [SE06] is an extension to IKEv2 [C.K05] that provides mechanisms so that mobile clients with VPN (Virtual Private Network) connectivity do not need to tear down the existing security association during their layer 3 hand-off. In base IKEv2 protocol, the IKE SAs and tunnel mode IPSec security associations are created implicitly between the IP addresses that are used when IKE_{SA} is established. When the mobile moves, its IP address changes giving rise to the need for a new IKE process and new security association between the mobile’s new IP address and the VPN gateway. This process results in suboptimal operation and media interruption. However, MOBIKE allows
the IP addresses associated with IKEv2 and tunnel mode IPSec Security Associations to modify by initiating additional signaling message such as UPDATE_SA_ADDRESS. As a practical application, a mobile Virtual Private Network (VPN) client could use MOBIKE to keep the connection with the VPN gateway active while the mobile itself changes its address due to change in network point of attachment. Similarly, a multi-homed host could use MOBIKE to move the traffic to a different interface if, for instance, the one currently being used stops working. MOBIKE is probably the most preferred mobility protocol to take care of mobility for the clients in mobile VPN environment, where the mobile does not need any additional home agent. In Chapter 5, I have proposed optimization techniques to improve the handoff delay performance for a mobile in a VPN environment.

Table 2.3 shows a qualitative comparison of some of the available IP-based mobility management protocols in the wireless Internet. Only network layer, transport layer and application layer mobility protocols are cited in this comparison. The mobility protocols with an “*” next to it are the candidate protocols that I have experimented with as part of my thesis. I briefly define the metrics against which these protocols are compared. Intra-domain encapsulation involves extra level of encapsulation due to additional tunneling in the network. Inter-domain encapsulation involves tunneling when the mobile moves between the domains. End system changes means whether additional software is needed to make that specific mobility protocol to work. Fast handoff support means whether the handoff delay is reduced.
Table 2.3: Qualitative comparison of IP-based mobility protocols

<table>
<thead>
<tr>
<th>Mobility Protocol</th>
<th>Intra-domain encapsulation</th>
<th>Inter-domain encapsulation</th>
<th>End system changes</th>
<th>Triangle routing</th>
<th>Network change</th>
<th>Fast-handoff technique</th>
<th>Mobility layer</th>
<th>Mobility type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IPv4*</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
<td>--</td>
<td>Network</td>
<td>Macro</td>
</tr>
<tr>
<td>MIPv6*</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>Network</td>
<td>Macro</td>
</tr>
<tr>
<td>MIP-RO</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Network</td>
<td>Macro</td>
</tr>
<tr>
<td>MIP with FA assisted</td>
<td>X</td>
<td>X</td>
<td>--</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Network</td>
<td>Macro</td>
</tr>
<tr>
<td>IDMP*</td>
<td>X</td>
<td>X</td>
<td>--</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Network</td>
<td>Micro</td>
</tr>
<tr>
<td>MIP with LR*</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>Network</td>
<td>Macro</td>
</tr>
<tr>
<td>Cellular IP*</td>
<td>--</td>
<td>X</td>
<td>--</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Network</td>
<td>Micro</td>
</tr>
<tr>
<td>HAWAII</td>
<td>--</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Network</td>
<td>Micro</td>
</tr>
<tr>
<td>MSOCKS</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Transport</td>
<td>Macro</td>
</tr>
<tr>
<td>TCP-Migrate</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Transport</td>
<td>Macro</td>
</tr>
<tr>
<td>SIP*</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>X</td>
<td>Application</td>
<td>Macro</td>
</tr>
</tbody>
</table>

"X" - Yes "-" - No
2.5.7 Multicast mobility

IP multicasting allows IP packets to be delivered from a single source to a group of receivers that are part of the same multicast group using multicast routing protocol. Multicast packets are generally routed along a single shared tree or multiple source-based spanning trees for efficient content distribution. There are several proposed schemes to provide native IP multicast routing over a wide area network such as PIM [DEF+94], MOSPF [Moy93], DVMRP [WPD88], CBT [BFC93], BGMP [Tha04]. Traditional multicasting techniques do not handle large number of distinct multicast groups and do not provide means to handle multicast when some routers may not be multicast capable, while others are. Recently, there are however alternate techniques being developed, such as Source Specific Multicast (SSM) [HC06], UMTP (UDP Multicast Tunneling Protocol) [Fin03] and AMT [TV+02] that provide multicast support for the non-multicast enabled networks while providing novel ways to support specific application such as content distribution. Quinn and Almeroth [QA01] explain many of the issues associated with multicast deployment over wide area networking. Security, scalability and interoperability are some of deployment issues associated with multicast. However, I focus on mobile receiver issues as part of my thesis. Multicast group “join” and “leave” latencies are some of those deployment issues related to multicast mobility. Joining and leaving the multicast groups are usually handled through IGMP [Fen97], [CDK+02] although I have developed some alternative application layer techniques as part of my thesis and have described those in Chapter 8. In this section, I only highlight the mobility aspects of multicast traffic.

Unlike unicast traffic, multicast traffic distribution is receiver driven, where the mobile receiver makes a request to its access router to join a specific multicast group and router joins the upstream multicast tree. The mobile receiver periodically exchanges its group information with the router using Internet group management protocol (IGMP) [CDK+02] in IPv4 or multicast listener discovery (MLD) [ABO97] in IPv6 network. Mobile multicast introduces several challenges, namely general multicast routing issues, mobile receiver is-
CHAPTER 2. INTRODUCTION TO MOBILITY MANAGEMENT FOR MULTIMEDIA

issues, mobile source issues and deployment issues. Movement of a sender and receiver can introduce encapsulation and decapsulation of packets and routing state maintenance due to dynamic creation of multicast tree. Multicast receiver issues include join latency, packet loss, packet duplication, packet out-of-order and leave latency that affect the performance of real-time traffic. Reverse path forwarding (RPF), packet loss, multicast scoping, and source discovery are some of the issues associated with multicast delivery when the source is mobile. I describe some of the multicast mobility related work in this section.

Xylomenos and Polyzos [XP97], Varshey and Chatterji [VC99], and Acharya et al. [ABN95] describe many of the architectural issues associated with mobility support for multicast traffic. Romdhani [RKL’04] et al. categorize mobile multicasting primarily into four categories: home subscription-based, remote subscription-based solutions, hybrid solutions, and non-IP multicast-based solutions. I describe two of these approaches and some of mobile multicast protocols in each category.

2.5.7.1 Home subscription-based approach

Home subscription-based or bi-directional tunneling approach depends on home network and associated mobile IP entities such as home agent, foreign agent and uses a multicast router located in the home network. It puts the multicasting burden on the Home Agent (HA) by creating tunnels between the HA and the mobile or FA. However, tunneling multiple multicast packets to the foreign network is inefficient.

Figure 2.21 illustrates the home subscription-based architecture. Figure 2.22 describes the associated flow. In this architecture, the HA is multicast enabled and is responsible for periodically forwarding multicast group membership control messages to the mobile while it is away. To join a specific multicast group, the mobile node establishes a bi-directional tunnel with its HA and tunnels its membership message (e.g., IGMP) to the HA. It uses the same tunnel header used for routing unicast packets between the MN and HA. When the HA receives the join request it decapsulates and forwards it to the local multicast router on
the home link. The local multicast router (MR) intercepts this membership message and sends the join message to the upstream router on the home network. Once the multicast branch is established, the HA forwards the multicast traffic destined to the mobile over the tunnel. Thus, the HA acts like a multicast proxy for the mobile. When the mobile moves to the new foreign subnet, it does not need to join the multicast tree again, as the HA has already joined the multicast tree on behalf of the mobile. Although this method has the advantage that the mobile does not need to rejoin the multicast tree when the mobile moves from one network to another. This scheme suffers from triangle routing and tunnel overhead resulting in join latency. The HA needs to establish per-MN-tunnel to forward the multicast traffic. HA is also the central point of failure in this scheme.

![Diagram of multicast mobility: Home subscription-based](image)

*Figure 2.21: Multicast mobility: Home subscription-based*

There are several mobile multicast protocols that belong to home subscription-based category and improve upon some of these drawbacks associated with basic home subscription-based approaches. I highlight some of these multicast protocols here. Mobile Multicast (MoM) [WHMB98] proposes to reduce the explosion problem in bi-directional tunneling
by electing one designated HA called designated multicast service provider (DMSP). MoM protocol uses single tunnel between the HA and the FA instead of using per-MN tunnel. FA in return uses link-level multicast to complete the delivery. Range-based MOM [LW00] takes MoM approach one step further and elects a multicast agent close to FA to tunnel multicast packets to the foreign network. MPDSR (Multicast Protocol With Dynamic Service Range) [YP01] enhances RBMOM protocol in the Mobile IPv6 environment. The main goal of this approach is to determine an optimal service range and avoid service disruption. It does so by introducing two other network elements such as core source node (CSN) and boundary foreign agent (BFA). In order to eliminate possible disruption triggered by the handover procedure, the mobile receiver pre-joins the multicast group prior to the handover.
2.5.7.2 Remote subscription-based approach

Remote subscription-based approach takes the burden off the home agent and does eliminate tunneling and avoids the duplication of multicast packets being tunneled to foreign networks. However, this requires that after each handoff, the user must rejoin a multicast group. In addition, the multicast trees used to route multicast packets will be updated after every handoff to track the multicast group members.

As shown in Figure 2.23, unlike home subscription-based approach, in remote subscription-based approach, the mobile receiver joins the multicast group via a local multicast router (MR1) on the foreign network. The mobile sends its membership report message to the local multicast router (MR1) in the visited network, where the mobile currently resides. The mobile node and the multicast router take care of group management using IGMP for IPv4 and MLD for IPv6. The local multicast router intercepts this message and joins the desired multicast group by joining the upstream routers. Figure 2.24 shows the flow associated with the remote subscription-based approach. Since this approach does not depend upon the home network or any of the home network elements, after the handover, the mobile obtains a new care-of-address (CoA) and sends a new membership report to the multicast router (MR2) in the new network. While MR2 is in the process of joining the upstream multicast tree, the router in the previous network MR1 leaves the upstream multicast tree if there is no other receiver tuned to the multicast group. This is called leave latency as defined in Appendix C. After the mobile node moves to the new network, it does not have to wait to complete mobile IP registration before it is able to join the multicast tree in the new network. Thus, the join latency in case of remote subscription-based approach is less than that of home subscription-based approach.

There are several remote subscription-based mobile multicast protocols available that try to reduce the join latency and tunnel convergence problem. Some of these protocols include: remote subscription-based Mobile IP, Mobicast [TP00b], hierarchical SSM [KHHK01], remote subscription with multicast agent, pe-registration with mobility sup-
Figure 2.23: Multicast mobility: Remote subscription-based

port agent (MSA) [Wu99], multicast agent protocol (MMA) [SSK00], timer-based mobile multicast protocol (TBMOM) [Par02]. Mobility support agent (MSA) proposes a solution based on pre-registration to reduce the join latency and associated packet loss. Multicast agent protocol (MMA) uses a forwarding technique between the foreign networks instead of pre-registration method. This reduces the packet loss due to join latency. Timer-based mobile multicast protocol (TBMOM) selects a foreign multicast agent (FMA) to store the membership information of the mobile members in the foreign network. TBMOM protocol reduces multicast packet loss since unicast tunnels can be set up between FMAs of different foreign networks. Hierarchical SSM-based (Single Source Multicast) and Mobicast-based approaches provide hierarchical mobility management approaches by introducing hierarchical mobility agents such as border gateway router (BGR) and domain foreign agent (DFA), respectively. Both of these approaches split the multicast path into two level hierarchy and thus limits the multicast traffic distribution in the lower hierarchy whenever there is a movement within a hierarchy.
I provide a summary of these two types of multicast protocols and their qualitative comparisons with respect to some of the performance parameters such as “JOIN” latency, point of failure, tunnel convergence, and hierarchical mobility properties in tables 2.4 and 2.5, respectively. Tunnel convergence problem arises when a number of mobiles belonging to different home networks are anchored at a specific foreign agent. If all those mobiles are subscribed to a specific multicast group, then each HA responsible for the corresponding mobile tunnels the multicast packet to the FA. This will result in duplicate packets for the same multicast group. Number of duplicate packets will increase as the number of mobile hosts subscribed to the same multicast group increases. Thus, it is desirable to reduce the problem due to tunnel convergence. I briefly explain the comparison of “JOIN latency” for both home subscription- and remote subscription-based protocols. For example, average join latency of home subscription-based approach is independent of multicast group size, as each mobile receives multicast datagram from its home agent independently. Average join latency for MOM is moderate and is similar to home subscription-based approach.
However, disruption in MOM protocol decreases with the increasing in group density due to the sharing of DMSP by the visiting group members of the same group in the same foreign network. Compared with home subscription and MOM, remote subscription (RS) approach and multicast agent approach provide less disruption since access to the multicast backbone is much closer. However, multicast agents have a lower disruption than remote subscription when multicast group size is small. As the group size increases, remote subscription method does better than the multicast agent approach. However, at a very high group density, both the approaches offer same performance as the probability of finding other group members in a new IP subnet approaches 1. Mobility support agent (MSA) approach reduces the “join” latency to a great extent due to its pre-registration ability before the handoff.

<table>
<thead>
<tr>
<th>Protocols</th>
<th>JOIN latency</th>
<th>Point of failure</th>
<th>Hierarchical mobility</th>
<th>Tunnel convergence</th>
<th>Key differentiator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IP Home Subscription</td>
<td>Remains same with group size</td>
<td>Home Agent</td>
<td>No</td>
<td>One tunnel per MN</td>
<td>-</td>
</tr>
<tr>
<td>MOM</td>
<td>Reduces with increase in group size</td>
<td>Designated Multicast Service Provider</td>
<td>No</td>
<td>One tunnel for all MNs</td>
<td>Introduces DMSP</td>
</tr>
<tr>
<td>RBMOM</td>
<td>Smaller than MOM</td>
<td>Designated Multicast Service Provider</td>
<td>No</td>
<td>One tunnel for all MNs</td>
<td>Introduces service range</td>
</tr>
<tr>
<td>MPDSR</td>
<td>Least handoff delay</td>
<td>BFA</td>
<td>No</td>
<td>One tunnel for all MNs</td>
<td>Reduces tunnel length</td>
</tr>
</tbody>
</table>

Multicast mobility protocols in their current form suffer from performance issues due to tunnel overhead, join, and leave latency. Thus, optimization techniques can be applied
## Chapter 2. Introduction to Mobility Management for Multimedia

Table 2.5: Multicast mobility - Remote subscription

<table>
<thead>
<tr>
<th>Mobility Protocols</th>
<th>JOIN latency</th>
<th>Point of failure</th>
<th>Hierarchical mobility</th>
<th>Tunnel convergence</th>
<th>Differentiator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote subscription</td>
<td>Less than home subscription</td>
<td>FA or DHCP</td>
<td>No</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Multicast Agent</td>
<td>Same as RS</td>
<td>MA</td>
<td>No</td>
<td>Tunnel between FA and MA</td>
<td>Uses coordinator MA between FAs</td>
</tr>
<tr>
<td>MSA</td>
<td>Less than RS (pre-registration)</td>
<td>MSA</td>
<td>No</td>
<td>No tunnel</td>
<td>Pre-registration</td>
</tr>
<tr>
<td>MMA</td>
<td>Less packet loss due to forwarding</td>
<td>MA and MF</td>
<td>No</td>
<td>One IP tunnel (MA and MF)</td>
<td>Forwarding technique reduces packet loss</td>
</tr>
<tr>
<td>TBMOM</td>
<td>Less than RS</td>
<td>DMSP</td>
<td>No</td>
<td>Does not use IP tunnel</td>
<td>Uses hybrid forwarding approach</td>
</tr>
<tr>
<td>Hierarchical SSM</td>
<td>Less delay than RS</td>
<td>BGR</td>
<td>Yes</td>
<td>IP tunnel between source and BGR</td>
<td>Reduces JOIN latency for micro mobility</td>
</tr>
<tr>
<td>Mobicast</td>
<td>Less packet loss buffering</td>
<td>DFA</td>
<td>yes</td>
<td>Does not use IP tunnel</td>
<td>Buffering solves packet loss</td>
</tr>
</tbody>
</table>

To the existing multicast mobility protocols to reduce the join latency and leave latency. In Chapter 8, I present some of the optimization techniques that I have developed for multicast stream delivery that reduces the join latency in a hierarchical multicast environment.