

Gap Analysis and Deployment Architectures for 3GPP2 MMD Networks

Tsunehiko Chiba, Hidetoshi Yokota, Akira Idoue, KDDI R&D Laboratories, Inc.
Ashutosh Dutta, Subir Das, Fuchun Joseph Lin, Telcordia Technologies, Inc.

In recent years, mobile and fixed operators are offering broadband access to the Internet by either adding wireless access to their existing core networks (e.g., CDMA 1xEVDO, GPRS, and WiFi) or continuing to evolve existing broadband infrastructures (e.g., DSL, Cable and FTTx). Consequently, IMS and MMD architectures as defined by 3GPP and 3GPP2 are becoming increasingly important to network operators, in particular, for offering ubiquitous service and seamless mobility to end users, along with management and control. Although 3GPP2 MMD architecture has adopted most of the 3GPP IMS core functionalities and interfaces, it added several other unique functional components and interfaces to fit the specific need of CDMA2000 access networks. Therefore, it is important to understand the gaps and issues associated with 3GPP2 MMD architecture with regard to 3GPP IMS architecture. In this paper, we first analyze the functional and protocol level differences between IMS and MMD architectures. Based on our analysis, we then articulate several deployment architectural alternatives for 3GPP2 MMD networks and finally analyze the pros and cons of each architecture.

1. Introduction

To provide ubiquitous seamless service across multiple heterogeneous access technologies, network providers need an infrastructure that supports and controls a variety of network functionalities, such as quality of service, routing, security, and mobility, among others. The 3GPP (Third Generation Partnership Project) and 3GPP2 (Third Generation Partnership Project 2) developed the core IP multimedia architectures called IMS (IP Multimedia Subsystem) and MMD (Multimedia Domain), respectively, to meet these requirements. These architectures are viewed by the industry as the common infrastructure for the evolution of FMC (Fixed Mobile Convergence) and NGN (Next Generation Networks). However, 3GPP2 MMD has some differences from 3GPP IMS architecture, protocols, and interfaces. Therefore, it is important for 3GPP2 operators who are planning MMD deployment to identify the key differences and to address those deployment issues unique to 3GPP2 access and core networks. In this paper, we first address the differences in architecture,

protocol interactions, and communications between certain functional components in the MMD and IMS framework. Then, we propose three alternative deployment architectures that can be realized over an MMD network. An initial analysis evaluates these architectures using a common set of metrics. The pros and cons of each of these architectures are examined. This analysis also includes several important functionalities, such as configuration, P-CSCF discovery in the visited network, and interaction with the Mobile IP HA (Home Agent).

The rest of the paper is organized as follows: Section 2 provides gap analysis of IMS and MMD architectures at a functional level. Section 3 provides alternative deployment architectures for a 3GPP2 network by considering the placement of proxy servers. Section 4 provides a comparison and discussion of alternative architectures, and finally Section 5 concludes the paper.

Table 1 provides a list of abbreviations used in this paper.

Table 1 Abbreviation List

AAA	Authentication, Authorization, and Accounting
AP	Access Point
AS	Application Server
CCF	Charging Collection Function
CSCF	Call Session Control Function
HSS	Home Subscriber Server
I-CSCF	Interrogation CSCF
IM-SSF	IP Multimedia - Services Switching Function
ISC	IMS Service Control
MGW	Media Gateway
MGCF	Media Gateway Control Function
MRFC	Media Resource Function Controller
MRFP	Media Resource Function Processor
OSA	Open Service Access
OSF	Operation Systems Function
P-CSCF	Proxy CSCF
PDF	Policy Decision Function
PDIF	Packet Data Interworking Function
PSTN	Public Switched Telephone Network
SCS	Service Capability Server
S-CSCF	Serving CSCF
SLF	Subscriber Location Function
UE	User Equipment

2. Gap analysis of 3GPP IMS and 3GPP2 MMD Architectures

2.1 Architectural and Protocol Differences

Figure 1 presents the configuration of IMS and MMD entities as described in [1] and [2]. We identified a few entities that are different in the MMD architecture compared with the IMS architecture. These differences are represented by the shaded regions. In MMD, functional entities are interconnected by both reference points (RPs) and interfaces. Interfaces are represented as upper case letters with a smaller case suffix (e.g. Cx, Gi, etc.). As defined in [2], reference points are conceptual points that divide two groups of functions. These reference points are labeled in boldface in Figure 1. The major architectural differences are the authentication sub-system and mobility management. In IMS, HSS is an independent entity and has several functional components, including mobility management, security, location information, and identity handling. On the other hand, MMD split the functionalities into two distinct components, specifically a AAA server and a Database. While AAA and Database in MMD are functionally equivalent to an IMS HSS, there are subtle differences. Location information also employs the position server and proxy concepts in MMD. The access gateway can directly access the position server and request registration for a particular user, whereas in IMS, location information is only available in the HSS and can be accessed by either I-CSCF or S-CSCF. Mobility management in MMD is handled by the Mobile IP, an IETF (Internet Engineering Task Force) Standard, while IMS mobility management is supported by GTP (Generic Tunneling Protocol), the GPRS (General Packet Radio Service) mobility protocol.

There are differences in terms of interfaces, such as Si, Dx, Go, Gq, Rx, and Gx, although MMD architecture adopted most of the IMS interfaces. Since MMD does not support a CAMEL application, MMD does not have a Si interface, which is the interface between HSS and CAMEL. The Dx interface in IMS lies between SLF and CSCF, which is used to find the location of HSS when there are multiple HSS for users.

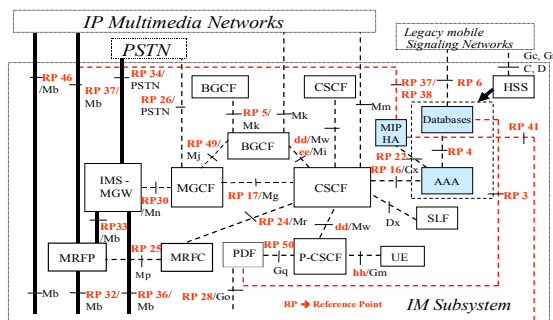


Figure 1 Configuration of IMS and MMD Entities

Figure 2 depicts the functional architecture that provides support for service provisioning for IMS as defined in [3]. The architecture and functional entity differences are represented by the shaded regions. The major difference in this architecture is IM-SSF, which is a particular type of application server to host the CAMEL network features, such as trigger detection points, and to inter-work with CAP (CAMEL Application Protocol). As described above, this is a 3GPP-specific feature that supports legacy services only with IM-SSF and the CAP interface. However, MMD specifies additional interactions, such as between AAA and position servers. While no standard interfaces are defined for this purpose, the type of interactions may be different for location-based services, for example.

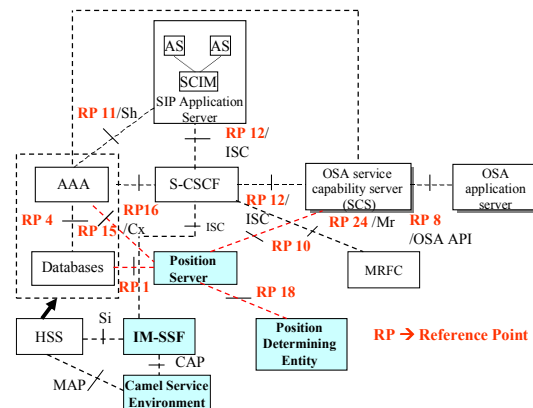


Figure 2 Functional Architecture for Services in IMS

2.2 Gap Analysis at the Functional Level

This subsection analyzes the similarities and differences at a functional level between IMS and MMD as identified in the previous subsection. We also address the key technical issues and challenges that a CDMA2000 operator needs to consider before deploying MMD networks.

• Signaling

SIP (Session Initiation Protocol) [4] and SDP (Session Description Protocol) [5] are used for call control in MMD. 3GPP2 documents [6] and [7] describe the signaling and session handling procedures that are based on 3GPP documents [8] and [9].

• IP version

While 3GPP IMS was designed for IPv6 only, 3GPP2 provides support for IPv4 as well as IPv6.

• QoS (Quality of Service)

Due to the fact that access networks are different for 3GPP and 3GPP2 networks, the specific protocols and parameters for providing QoS are also different. For example, PDF in IMS connects to GGSN (Gateway GPRS Support Node), whereas PDF in MMD connects to PDSN (Packet Data Serving Node). In IMS, GGSN is always an anchor, and therefore, the Go interface that connects the PDF and GGSN is always fixed. On the other hand, in MMD, PDF and PDSN may be in the visited

network. In such cases, local policies can be applicable. In addition, the Go interface in IMS is based on COPS (Common Open Policy Service) or Diameter [10], but in MMD, it is based on RADIUS (Remote Authentication Dial In User Service) [11] or Diameter.

- **Mobility**

Both 3GPP and 3GPP2 provide IP mobility, but the protocol and interfaces are different. 3GPP2 uses Mobile IP as the mobility management protocol and requires every subnet to have an FA (Foreign Agent) in the case of IPv4. It integrates the mobility management with AAA while MNs (Mobile Node) roam to a different AAA domain. On the other hand, in 3GPP, mobility is handled by the 3GPP-specific mobility protocol, the GTP that is used in GPRS Networks. The 3GPP specification does not integrate AAA with mobility.

- **Security**

3GPP uses an HSS for authentication and authorization, but 3GPP2 uses two entities, a AAA and a Database (which are together sometimes referred to as HSS). Although 3GPP2 interfaces are the same as a 3GPP interface, its user identity storage may be different in the HSS.

- **Billing**

3GPP has defined a separate entity called CCF for charge data collection, while 3GPP2 uses the AAA framework to perform the same function.

- **Placement of P-CSCF**

In IMS, the P-CSCF must be located within the same network as GGSN, i.e., P-CSCF and GGSN are bound. Therefore, IMS does not currently support a configuration where the GGSN is located in the home network while the P-CSCF is in the visited network. However, in MMD, the location of P-CSCF can be anywhere, either in the visited network or in the home network.

3. Alternative Architectures for MMD Networks

From the above analysis, one of the major differences identified for IMS and MMD is that in MMD, an operator can place the P-CSCF anywhere in the network based on specific requirements and ease of management. However, the placement of P-CSCF may affect the operation of the network in terms of signal load, scalability, operator roaming, and handover optimization. This section presents some candidate deployment architectures and discusses the associated pros and cons of these architectures. For a reliable operation, there are several key issues that need to be considered for each of these architectures, for example, SA (Security Association) creation, tunnel management, P-CSCF discovery, handover optimization, scalability, and management.

- **SA creation**

IMS requires that every MN must establish an IPsec SA with the P-CSCF before a session begins. While the MN is in session and moving to a new P-CSCF (assuming every visited network has one P-CSCF), it needs to create a new SA. This would not allow carrying the same session seamlessly.

- **Tunnel management**

Mobile IP requires establishing a tunnel between the MN and the FA or the FA and the HA in IPv4 and between the MN and the HA in IPv6. MMD also requires an IPsec tunnel between the MN and the P-CSCF. An MN, therefore, needs to maintain multiple tunnels, which may be an issue for smaller handheld devices.

- **P-CSCF Discovery**

As the MN moves between a home and a visited network or between visited networks, it needs to communicate via its first outbound SIP proxy, which happens to be the P-CSCF in the IMS and MMD architectures. P-CSCF is the first logical entity in the SIP signaling plane. There are several ways that P-CSCF discovery can take place based on the type of architecture, for example, static provisioning, DHCP INFORM, DNS SRV, and router advertisement.

- **Handover optimization**

Each P-CSCF retains the state of certain aspects of an ongoing call. This call state includes certain metrics of the ongoing call, such as media information, QoS, and call data record. Thus, when the MN moves to a new network and needs to communicate with the new SIP server, the call state needs to be transferred [12] as quickly as possible so as to maintain the same call quality in the new network. In addition, all the SIP-related signaling needs to go through the P-CSCF before it reaches to I-CSCF and S-CSCF. The P-CSCF discovery described above is also in the critical path of handover optimization, although a Mobile IP binding update message does not need to go through the P-CSCF. Efficient and timely discovery of P-CSCF and efficient context transfer of states between P-CSCFs can help optimize handover performance.

- **Scalability and management**

Since a P-CSCF is in the call path during handover, the placement of the P-CSCF may give rise to a number of signaling messages over wireless access networks. Wireless resources are scarce; thus, the total number of signaling messages needs to be reduced. When an MN is inside a visited network, it can move between several sub-networks within the visited network, where each sub-network can be a subnet associated with a specific PDSN. Primarily, P-CSCF can be optionally placed in three different parts of the logical network associated with the MN, the home network, the visited network, or the sub-networks of the visited network. Each of these scenarios gives rise to a different set of challenges and deployment strategies. This placement scenario can be compared with a typical case of a completely distributed vs. a centralized architecture. For example, in a centralized architecture, an MN has to rely on some central authority to be able to obtain the address of a P-CSCF, whereas in a completely distributed architecture, an MN needs to rely on some visited network components, such as a DHCP server, to obtain the P-CSCF address. A completely distributed architecture reduces the dependence on a central server and also reduces the failure rate at the expense of more signaling

message exchanges and discovery costs at the edge of the network. A middle ground could be set where these servers can be placed within a visited network; thus, as long as the MN moves within the visited network, it does not need to change the P-CSCF unless there is a need for load balancing. We describe each of these architectural alternatives in the following subsections.

3.1 Alternative Architecture I

This subsection describes a model where P-CSCFs are placed in the home network. There can be more than one P-CSCF in the home network to balance the load. We assume that a provider comprises one or more domains, and each domain comprises one or more networks. These networks can either be home networks or visited networks with reference to MNs. A network can further be divided into subnets (a.k.a. sub-networks), where each network may consist of one or more subnets. Thus a visited network may consist of several sub-networks. Basically, a domain can be distinguished from the others in terms of a DNS or a AAA entity. On the other hand, a provider’s domain can also have multiple networks, each with different access mechanisms, such as CDMA2000, WiFi, and WiMAX.

Figure 3 illustrates the architecture and the functional components. When an MN is initially configured, it is configured with a specific P-CSCF from a set of P-CSCFs through a DHCP procedure within the home network. As MN makes the transition to a new visited network, it re-registers with the HA using a new FA care-of-address for IPv4 networks or a new co-located care-of-address for IPv6 networks. In this architecture, the DHCP server assigns a P-CSCF address to the MN.

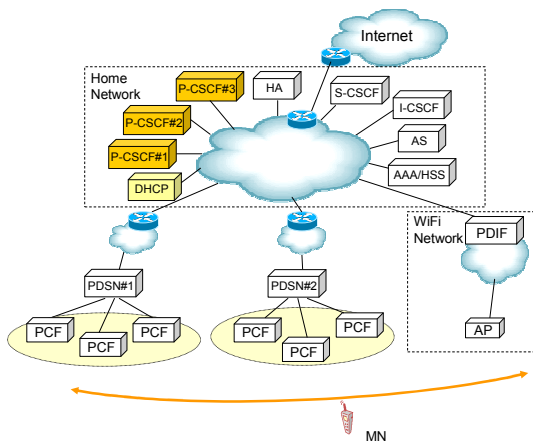


Figure 3 Architecture I: P-CSCFs in the home network

Figure 4 shows a specific IPv6-based call flow related to the architecture scenario presented in Figure 3. For brevity, an I-CSCF is not shown in the flow diagram. The MN sends the SIP registration message to the assigned P-CSCF and a Mobile IP binding update message to the HA to update its co-located

care-of-address. This way, the signal and media are all routed through the HA in both directions if reverse tunneling is set up. This is an example where the MN first makes a transition from one visited network to another visited network involving the same radio access technology (such as CDMA2000) and then makes an inter-access handover, thereby moving from one access technology to another one (such as from CDMA2000 to WiFi). When MN performs an inter-PDSN handover within the same CDMA2000 network, it receives the same P-CSCF address through a DHCP Ack message. At this time, the DHCP server has the option of sending a different P-CSCF address, which is purely determined by the policy of the home network and the load for serving the P-CSCF. In order to accomplish this, the DHCP server needs to keep track of the available P-CSCFs that can serve a specific MN. When the MN performs an inter-access handover, it performs another DHCP procedure. In this typical example of a heterogeneous handover, the MN can also receive the same P-CSCF address, since the P-CSCF is placed in the home network and the DHCP server can assign the previous P-CSCF address by keeping track of the P-CSCF of the MN.

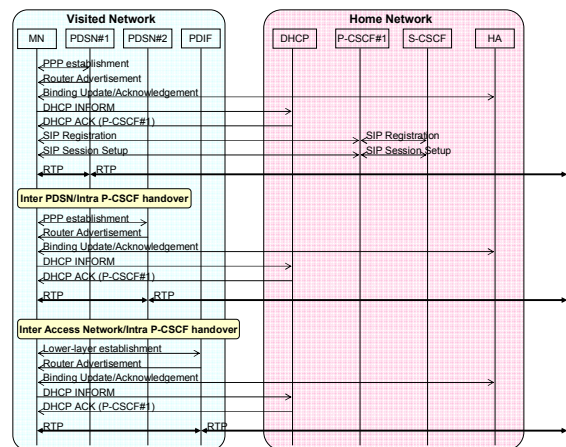


Figure 4 Call Flow for Architecture I

The pros and cons of Architecture I are described as follows.

- **Pros**
 1. Operator will have complete control of the operation of P-CSCF because it is in the home domain.
 2. Security between P-CSCF and HA will be trivial because both are situated in the home network.
 3. Call state transfer between P-CSCFs may not be needed.
 4. Since there may be more than one P-CSCF configured in the home network, the DHCP server can easily keep track of each P-CSCF load and assign an appropriate P-CSCF in order to avoid the failure of a specific server due to excessive signaling.
- **Cons**
 1. This architecture has a scalability issue due to its completely centralized architecture.

2. Since the P-CSCF discovery is based on the DHCP server, which is located in the home network, the failure probability of P-CSCF discovery and its discovery time is higher due to the long traversal of DHCP messages.

3.2 Alternative Architecture II

This model provides a small departure from the centralized model described in the previous subsection. This subsection describes a model where all P-CSCFs are placed in the visited network. We subdivide this scenario by splitting it into two subcategories based on the physical location of the P-CSCF.

3.2.1 Alternative Architecture Ia

This subsection considers an architectural scenario where the visited network includes all P-CSCFs, but these are not distributed at the PDSN level, including the possibility within a WiFi network. Figure 5 shows an example where P-CSCFs are located in the core of the visited network. In this architecture, a visited network consists of multiple subnets with a PDSN at each subnet level. There are no P-CSCFs and DHCP server in the home network. The home network has an S-CSCF and may have an I-CSCF. The visited network may have an I-CSCF as well.

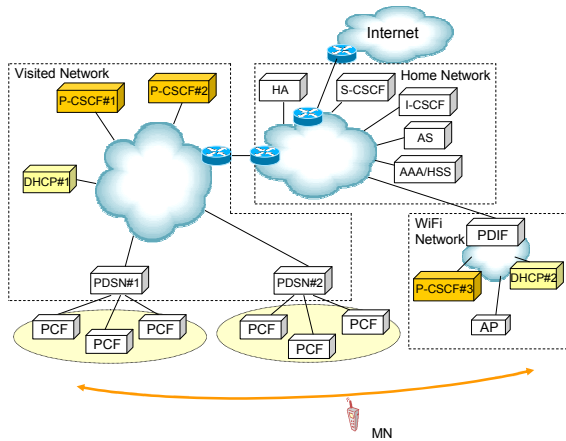


Figure 5 Architecture IIa: P-CSCFs in the visited core network

Figure 6 shows the IPv6-based call flow associated with handover as depicted in Figure 5. In this case, the MN initially moves from one sub-network to another within the same visited network (e.g., same access technology). Then, it makes another transition to another access technology (e.g., WiFi), thus resulting in an inter-access handover. During the intra-access handover, the MN may have the same P-CSCF assigned to it as it moves within the same visited network and thus discovering another P-CSCF is not needed. However, when the MN executes a heterogeneous handover by moving to a new visited

network, it obtains a new P-CSCF address meant for that specific visited network. Therefore, a new P-CSCF discovery is mandated only when the MN moves between two different visited networks.

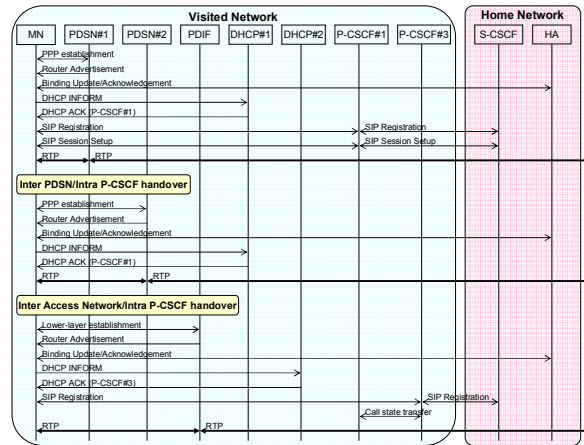


Figure 6 Call Flow for Architecture IIa

The pros and cons of Architecture IIa are described as follows.

- **Pros**
 1. This architecture provides flexibility in terms of management and control since it is semi-distributed. It also overcomes the drawbacks of a centralized control system.
 2. This architecture alleviates the need for signaling traversal to the home network for P-CSCF discovery.
- **Cons**
 1. Efficient call state transfer between P-CSCFs is needed when the MN moves to another visited network.
 2. SIP signaling has to traverse to the HA before being directed to P-CSCF because the P-CSCF is not in the home network. Similarly, the SIP response message, such as 200 OK, will need to go via the HA before being directed to the MN instead of P-CSCF sending the signal directly to the MN.

3.2.2 Alternative Architecture IIb

This subsection considers a completely distributed scenario where each visited network, potentially including a WiFi access network, can have a P-CSCF within it, and S-CSCF and I-CSCF stay in the home network. Figure 7 shows an example of such a distributed architecture. In this architecture, there is a P-CSCF associated with each PDSN in the visited network. Thus, the MN needs to discover a new P-CSCF when it moves between two different PDSNs or between a PDSN and a PDIF. The P-CSCF discovery procedure then becomes a local event and happens during the configuration phase at the sub-network level.

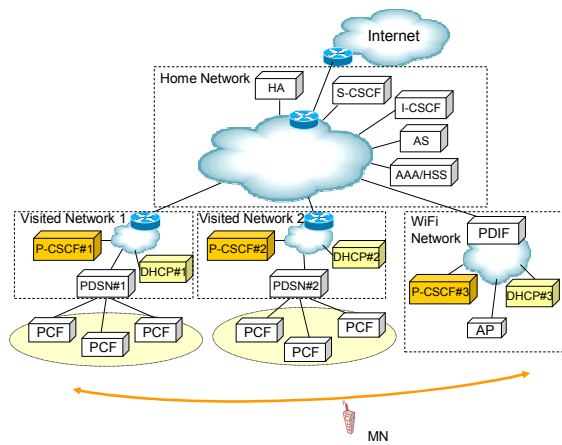


Figure 7 Architecture IIB: P-CSCF in each subnet

Figure 8 shows the IPv6-based call flow associated with Figure 7. In this case, there is a P-CSCF associated with each PDSN. Initially, the MN is connected to PDSN#1, which is in visited network 1. Since the MN is away from the home network, it obtains a new co-located care-of-address from the router advertisement and sends a Mobile IP binding update message accordingly. The MN also sends the SIP registration message to the assigned P-CSCF. In the case of handover, the MN completes a SIP registration procedure via the newly obtained P-CSCF, since the P-CSCF is different from the previously obtained P-CSCF. During the handover, it is absolutely essential that the call state is transferred from the old P-CSCF to the new P-CSCF to maintain the same level of call control. The first transition in the call flow example is between PDSNs, while the access technology remains the same. The second transition is between the PDSN and a new access technology such as WiFi. In each case, the Mobile IP binding update procedure and SIP registration are performed via the newly discovered P-CSCF. Any existing SA between the MN and the P-CSCF also needs to be re-established in the newly discovered P-CSCF, which may introduce a delay during handover.

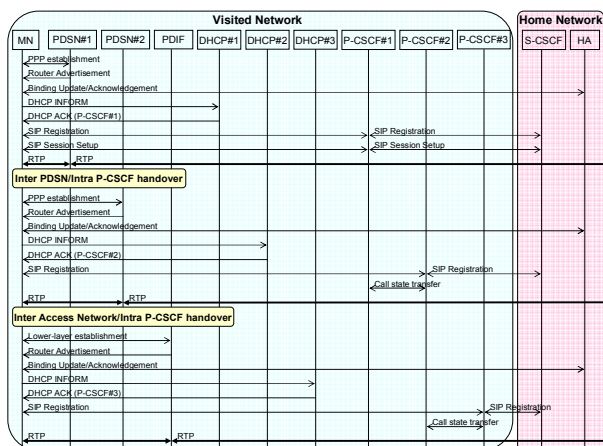


Figure 8 Call Flow for Architecture IIB

The pros and cons of Architecture IIB are described as follows.

- **Pros**
 1. This architecture is highly scalable and distributed.
 2. This architecture has less dependence on the home network.
 3. P-CSCF discovery in the local access network helps to reduce the signaling related to the discovery process.
 4. The visited network has ownership of the P-CSCF address.
- **Cons**
 1. Efficient call state transfer between P-CSCFs is needed when the MN moves to another PDSN or PDIF.
 2. Fast handoff between visited networks will give rise to a need for rapid P-CSCF discovery.
 3. SIP signaling has to traverse to the HA before being directed to P-CSCF because the P-CSCF is not in the home network. Similarly, the SIP response message, such as 200 OK, will need to go to the HA before being directed to the MN instead of P-CSCF sending the signal directly to the MN.

4. Comparative Analysis of Alternative Architectures

We provide a summary of the comparative analysis of these architectures. Architecture I provides the operator with complete control over the operation of P-CSCF, since it is inside the home network. Architecture IIA provides a middle ground between a completely centralized and distributed architecture and alleviates the need for signaling traversal to the home network for P-CSCF address discovery. Architecture IIB is highly scalable and distributed and is less dependent on the home network. P-CSCF discovery in the local access network helps to reduce the signaling overhead to the home network.

In Architecture I, managing the SA in P-CSCF and HA is relatively easier because both network entities are within the home network. Call state transfer between P-CSCFs is not generally necessary if the MN does not change its P-CSCF. Therefore, in Architecture I, a change of P-CSCF may be needed only when load balancing is needed. However, Architecture I shares the same disadvantage of being a centralized architecture. In most cases, the visited network can be far from the home network, thus P-CSCF discovery may take some time due to the traversal of signaling all the way to the home network.

If Mobile IP reverse tunneling is in effect, both Architecture IIA and IIB suffer from similar drawbacks, such as redundant routing, where SIP signaling has to go via the HA before being directed to the P-CSCF. Selection of a specific P-CSCF may be difficult for Architecture IIA since there is no dedicated P-CSCF for a specific PDSN. Although architecture IIB provides greater scalability and is distributed, fast handover between visited networks requires fast P-CSCF discovery in

this architecture. Handover optimization becomes more critical as the MN changes its P-CSCF frequently, giving rise to the need for call state transfer between adjacent P-CSCFs. It will also need an efficient method to transfer the call state between adjacent P-CSCFs.

Table 2 highlights some of the high level functional differences between the three architectures in terms of their impact on the operation in a real network. Based on the priority of a specific function, the operators can select a specific architecture.

Table 2 Comparative Analysis of Architectures

Architecture Type	Scalability	Local policy control	Operation complexity	P-CSCF handoff	SIP Signaling redundant path
I	Low	Low	Low	No	Short
Ila	Medium	Medium	Medium	Yes	Medium
Ilb	High	High	High	Yes	Long

5. Conclusions

This paper discusses and analyzes the functional and protocol level gaps between IMS and MMD architectures and provides a few alternative deployment architectures in MMD framework. We believe that the key findings of this analysis will be helpful for 3GPP2 operators in upgrading their next generation networks in order to provide IP multimedia services.

Acknowledgements

We would like to thank Prof. Henning Schulzrinne of Columbia University and the other members of Telcordia Technologies, Inc., for their valuable contribution.

References

- [1] 3GPP TS 23.002 V.6.10.0: "Technical Specification Group Services and System Aspects: Network Architecture," December 2005
- [2] 3GPP2 S.R0037-0 v3.0: "IP Network Architecture Model for cdma2000 Spread Spectrum Systems," September 2003
- [3] 3GPP TS 23.218 V6.3.0: "Technical Specification Group Core Network; IP Multimedia (IM) session handling; IM call model - Stage 2," March 2005
- [4] J. Rosenberg, et al, "SIP: Session Initiation Protocol," IETF RFC 3261, June 2002
- [5] M. Handley, et al, "SDP: Session Description Protocol," IETF RFC 2327, April 1998
- [6] 3GPP2 X.S0013-002-A v1.0: "All-IP Core Network Multimedia Domain; IP Multimedia Subsystem - Stage 2," November 2005
- [7] 3GPP2 X.S0013-004-A v1.0: "All-IP Core Network Multimedia Domain: IP Multimedia Call Control based on SIP and SDP - Stage 3," November 2005
- [8] 3GPP TS 23.228 V.6.11.0: "Technical Specification Group Core Network and Terminals; IP Multimedia Subsystem (IMS) - Stage 2," September 2005
- [9] 3GPP TS 24.229 V.6.8.0: "Technical Specification Group Core Network and Terminals; IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP) - Stage 3," September 2005

- [10] P. Calhoun, "Diameter Base Protocol," IETF RFC 3588, September 2003
- [11] C. Rigney et al, "Remote Authentication Dial In User Service (RADIUS)," IETF RFC 2865, June 2000
- [12] Kim Lynggaard Larsen et al, "Migration of the security association for fast SIP mobility within the IP multimedia subsystem," WPMC2006, September 2006