TECHNICAL REPORT

ISO/IEC TR 29181-4

First edition
2013-10-15

Information technology — Future Network — Problem statement and requirements —

Part 4: Mobility

Technologies de l'information — Réseaux du futur — Énoncé du problème et exigences —

Partie 4: Mobilité

Reference number
ISO/IEC TR 29181-4:2013(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide to publish a Technical Report. A Technical Report is entirely informative in nature and shall be subject to review every five years in the same manner as an International Standard.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 29181-4 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 6, *Telecommunications and information exchange between systems*.

ISO/IEC TR 29181 consists of the following parts, under the general title *Information technology — Future Network — Problem statement and requirements*:

— *Part 1: Overall aspects*

— *Part 2: Naming and addressing*

— *Part 3: Switching and routing*

— *Part 4: Mobility*

— *Part 5: Security*

— *Part 6: Media transport*

— *Part 7: Service composition*

Introduction

This part of ISO/IEC TR 29181 (Future Network: Problem Statement and Requirements) describes the problems of the current network and the requirements for Future Network in the mobility perspective. The general description on the problem statement and requirements for Future Network is given in ISO/IEC TR 29181-1. In addition, ISO/IEC TR 29181-4 establishes the problem statement and requirements for Future Network from the viewpoint of architecture and functionality for mobility support.

In general, the mobility issues can be classified into link-layer, network-layer, and transport/application layer mobility management. It is noted that the link-layer mobility issues have been addressed and well defined in the relevant SDOs, such as 3GPPs, IEEE 802, etc. The transport/application layer mobility issues are also associated with the particular transport/application protocols used by mobile nodes. On the other hand, the network layer mobility control issues are quite dependent on the network architecture. Accordingly, this part of ISO/IEC TR 29181 will focus on the mobility issues of Future Network in the network-layer perspective.

This part of ISO/IEC TR 29181 may be applicable to the overall design of Future Network architecture.

Information technology — Future Network — Problem statement and requirements —

Part 4: Mobility

1 Scope

This part of ISO/IEC TR 29181 describes the problem statements of current network and the requirements for Future Network in the mobility perspective. This part of ISO/IEC TR 29181 mainly specifies

— problems of the current network in mobile environment, and

— requirements for mobility support in Future Network.

In addition, this part of ISO/IEC TR 29181 gives information on

— existing mobility control schemes in the current network,

— examples of high-level mobility control architecture for Future Network,

— distributed mobility control in the Proxy Mobile IPv6 networks, and

— additional considerations for Future Network mobility.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC TR 29181 (all parts), *Information technology — Future Network — Problem statement and requirements*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1
Future Network
FN
network of the future which is made on clean-slate design approach as well as incremental design approach; it should provide futuristic capabilities and services beyond the limitations of the current network including the Internet

[SOURCE: ISO/IEC TR 29181-1]

3.2
Node Identifier
NID
globally unique identifier for a network node or host

[SOURCE: ISO/IEC TR 29181-2]

Note 1 to entry: Identifier (ID) is a generic term that is associated with various types of objects, whereas NID is used to represent a host in the network. In this part of ISO/IEC TR 29181, the term of ID is used with the meaning of NID.

3.3
Locator
LOC
IP address to connection mapping

[SOURCE: ISO/IEC TR 29181-2]

Note 1 to entry: In this part of ISO/IEC TR 29181, LOC is used to represent the location of a host in the network, which is also used for delivery of data packets in the network.

3.4
Mobility
ability for user to communicate or access the services, irrespective of changes of its location

[SOURCE: ITU-T Recommendation Q.1706/Y.2801]

4 Abbreviations

ACK	Acknowledgement
AR	Access Router
CN	Correspondent Node
FN	Future Network
GILL	Global Identifier and Local Locator
GW	Gateway
HA	Home Agent
HMI	Hierarchical MIP
ID	Identifier
IP	Internet Protocol
ISP	Internet Services Provider
LMC	Local Mobility Controller
LOC	Locator
MIP	Mobile IP
MN	Mobile Node
NID	Node ID
PMIP	Proxy Mobile IP

SCTP	Stream Control Transmission Protocol
TCP	Transmission Control Protocol
TR	Technical Report
UDP	User Datagram Protocol

5 General

5.1 Mobile environment in FN

With an explosive growth of the number of subscribers of 3G/4G cellular systems and also other wireless data systems such as WiFi and WiMAX, the mobile networks now become the key driver toward the Future Network (FN). The number of people who surf the network on their phones has doubled since 2006. In near future, there will likely be more mobile and wireless users than wired ones. In addition, a variety of new types of wireless access networks like ad hoc networks and sensor networks are emerging, and they will be the major access means to FN.

Figure 1 illustrates the network environment, in which the users or things in FN will benefit from a variety of access ways to the network anytime, anywhere, and through any interfaces. In particular, it is expected that 'mobile' users/things, rather 'fixed' ones, will become more dominant in FN. In this context, a crucial requirement for FN is to provide seamless services for the mobile users/things through the mobile-oriented FN.

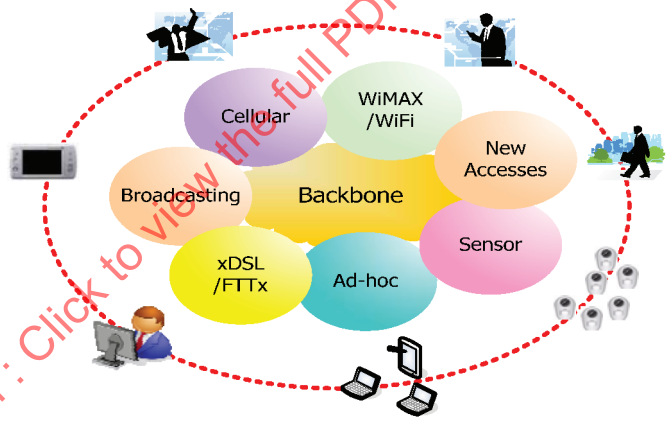

Figure 1 — Network environment in FN

With a recent trend of network convergence, it is expected that the all kinds of networks will be evolved or revolved toward a unified network, i.e. 'mobile-oriented convergence network' as shown in Figure 2, including computer or telecommunication networks.

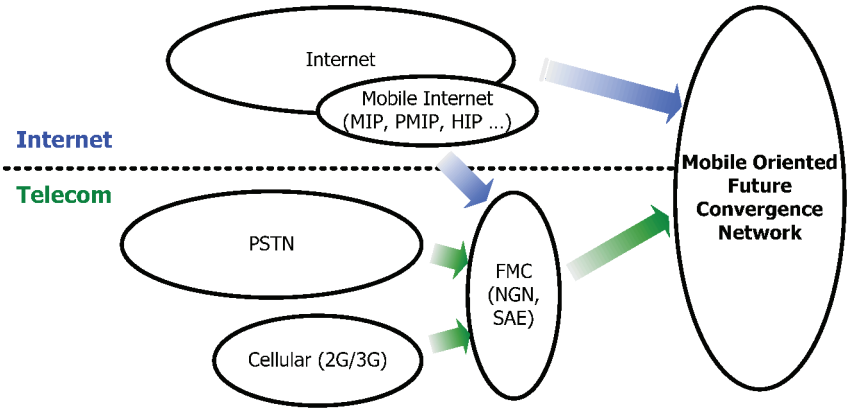

Figure 2 — Convergence of Internet and telecommunication in FN

Historically, the computer networks (i. e., Internet) and the telecommunication networks have been evolved until now, with quite different design philosophy and business purposes. From the perspective of the convergence trends, however, FN should be designed to make full use of the pros of both computer networks and telecommunication networks.

In addition, it is noted that the current cellular system is a very successful model in the wireless/mobile systems and provides a lot of desirable features to support the mobile environment. Therefore, it is recommended to readily exploit the useful features of cellular mobile systems in the design of FN architecture, to the extent possible.

5.2 Related works on mobility in FN

It is noted that the current network, such as Internet, was basically designed for fixed network environment, rather than for the mobile network environment. This has enforced Internet to add a lot extensional features to satisfy the requirements for mobile networks, as shown in the example of Mobile IP (MIP) that has been made in the IETF. The examples of mobility schemes for current network are described in the Annex A, which include the transport-layer and application-layer mobility schemes that have been so far proposed in the IETF. However, this patch-on approach seems to be just a temporal heuristic to the problems in the mobile environment, rather than an optimization approach to substantially solve the mobile-related issues.

Based on these observations, some activities already started to design the FN for mobile environment rather than fixed environment. A typical example is eMobility, which is a FP7 project of EU. eMobility envisions the third generation Internet as the wireless/mobile Internet with the name of 'Post-IP.' The 4WARD, another FP7 project for design of FN architecture, is also targeted to effectively support the mobile environment. The other European projects for mobility include Trilogy.

We also note a lot of Future Internet Design (FIND) projects, which are very closely related with wireless/mobile environments. Especially, the "Mobility First" project is recently proposed as a candidate approach for FN. This proposal is with the recognition that the network is changing very rapidly from fixed hosts to mobile devices, in which it is stated that a FN architecture should support mobile devices as 'first-class' users and also provide a variety of new applications efficiently, securely, and at a large scale. The Global Environment for Network Innovation (GENI), a representative testbed project for FN, also agrees that wireless/mobile will be the major access means for FN. We note that some design documents of GENI already covers the issues including ad hoc and sensor networks.

AKARI is a representative research activity on FN in Japan, which deals with the issues on the separation of Identifier (ID) and Locator (LOC) and the managed mesh network to support mobile environment. Especially, the ID-LOC split architecture covers the primary issues in mobile environment such as mobility and multi-homing, etc. On the other hand, Mobile Oriented Future Internet (MOFI) is a project of designing the architecture and protocols of FN in Korea.

In the perspective of standardization activities, the ITU-T has so far identified the mobility management requirements and frameworks for next-generation networks, which include ITU-T Q.1706, Q.1707, Q.1708, and Q.1709. It is noted that some preliminary works for Future Networks are also in progress in the ITU-T SG13.

From these observations, we can see that there is a crucial need to design the architecture of FN to more effectively support the mobile environment.

6 Problem statement of current network in mobile environment

6.1 Overloaded semantics of IP address

In the current IP-based networks, an IP address has overloaded semantics as Identifier (ID) and Locator (LOC). In mobile environment, however, the location of mobile host is likely to continue to change by movement. This means that the static allocation of LOC (IP address) to a host may become problematic in mobile networks. In the meantime, ID needs to be kept persistently (without change) to maintain an on-going sessions against movement of a host. Accordingly, ID and LOC should be separated to support the mobility in Future Network.

Another critical concern is that an IP address, as an ID, is allocated to a network interface of a host, rather the host itself. Accordingly, if a host has multiple interfaces, multiple IP addresses must be allocated to a single host. This may give serious inefficiency to a multi-homing host, since the same host has to use different IDs for communication. Therefore, ID needs to be allocated to a host itself rather than its network interface.

As for the allocation of IP address, it does not make sense to allocate IP address to a mobile host, since it may continue to move on. Accordingly, in mobile environments, it is suggested that an address or LOC should be allocated to a certain fixed node in the network, rather than the host itself.

6.2 Single common protocol for heterogeneous networks

It is expected that future mobile networks will consist of a variety of heterogeneous wireless networks. Such wireless networks are likely to have quite different characteristics, ranged from managed mobile networks to light-weight senor networks. On the other hand, the backbone network will be evolved to optical network with high bandwidth, which is quite different from wireless access networks. Accordingly, a single common IP protocol and/or addressing scheme of current Internet may not effectively support the FN with optical backbone and heterogeneous wireless access networks.

6.3 Integration of data delivery and control function

In most of current Internet protocols, data delivery and control function are integrated and implemented at the same devices, and the data and control traffics are routed along the same path, as shown in the IP and ICMP protocols. However, the control information for signalling is mission-critical and thus needs to be delivered more urgently and reliably, compared to usual data traffics. Thus, the control function needs to be separated from data traffics.

6.4 Centralized mobility control

Most of the current mobility control schemes in IP-based networks are based on a centralized mobility anchor, such as Home Agent (HA) of Mobile IP. This is because the existing mobile networks were originally designed as a hierarchical architecture to support circuit-based voice traffics. In the centralized control, however, the routing path through a centralized anchor tends to be longer, which results in non-optimal routes and performance degradation. Moreover, the centralized approach is vulnerable to a single point of failure or attack.

It is noted that an ever-increasing demand of mobile Internet traffics has enforced non-hierarchical or flat architecture on mobile networks, so as to provide data services more cost-effectively. Accordingly, we need to consider the distributed mobility control to support a flat architecture of future mobile networks.

7 Architectural requirements for mobility support in FN

7.1 Separation of identifier and locator

To avoid the overloaded semantics of IP address, ID should be separated from LOC (IP address). ID is used to identify a host itself in the network, whereas a LOC is used to represent the current location of a host in the network. ID is not an IP address, but a persistent ID. ID is given statically, not allocated dynamically. An IP address is used as LOC (e. g., IP address of an access router that the mobile host is attached to).

7.2 Support of heterogeneous access networks

Future Network should support a variety of heterogeneous access networks. To meet this requirement, the protocols and/or addressing schemes used for packet delivery may be different between access network and backbone network, and/or between access networks. In particular, the protocols for access networks may be designed by considering the heterogeneous wireless link characteristics, whereas the protocols for backbone networks may be designed to be as simple as possible by considering the optical networks.

7.3 Separation of mobility control function from user data delivery

In Future Network, to deal with mission-critical control information (e.g. for mobility control) effectively, it is recommended that the mobility control function should be separated from the user data plane. In particular, the mobility control function will be performed as a network-based scheme to enhance deployment, resource utilization and protocol performance.

7.4 Support of distributed mobility control

It is expected that the future mobile network will be evolved to a flat architecture, not a hierarchical structure. Accordingly, the Future Network should be designed to provide a distributed mobility control for flat network architecture. In the distributed mobility control, the route optimization can be intrinsically supported, and we can reduce unnecessary traffics flowing into the core network. This can also mitigate the problem of a single point of failure to a local network.

For information, the high-level architecture of mobility control in FN is discussed in the Annex B, and the use of Proxy MIP for distributed mobility control is discussed in the Annex C.

8 Functional requirements for mobility support in FN

8.1 Location management

To support the mobility, the Future Network should provide the location management function so as to keep track of the movement of a host in the network and also to locate the host for data delivery. The location management function is used for supporting the prospective 'incoming' session (or call) to the mobile user. The location management function includes the following sub-functions: location registration/update and location query/response (for user data transport) that may be performed with a service control function for call/session establishment.

The location registration and update functions are used to keep track of the current location of a host. When a host is attached to the network, it may register its current location with the network, possibly via an appropriate location database. When the host moves into the other network, the corresponding LOC will be updated. In the location registration and update function, the mapping information between IDs and LOCs will be managed and updated all the time.

The location query and response functions are used to locate a host for data delivery. The information of the current location of a host will be identified through the suitable location query and response operations. It is noted that the location query and response operations may be performed together with a relevant service control function.

8.2 Route optimization

To support the mobility, the Future Network should provide the route optimization function by which the two communicating hosts can exchange the data packets through a shorter path. This can also give performance benefits by reducing packet propagation delays, bandwidth consumption and congestions in the network.

8.3 Handover control

To support the mobility in Future Network, the handover control function should be provided to realize 'session continuity' for 'on-going' sessions of mobile hosts. To provide the seamless mobility or session continuity, the handover control functions will be performed to minimize the data loss and handover latency during handover.

In general, the handover control schemes can be divided by the protocol layer into the handover control in the link layer, the handover control in the network layer, and the handover control in the transport or application layer. Each of the handover schemes will be performed using the corresponding signalling between the entities associated with handover. The handover signalling will be based on the movement detection (in the link-layer and/or in the network layer). A different handover control protocol or scheme can be employed, depending on how to use the information on movement detection and/or how to perform the handover signalling.

It is noted that the link-layer handover issues have been addressed and well defined in the relevant SDOs, such as 3GPPs, IEEE 802, etc. The transport/application layer handover issues are also associated with the particular transport/application protocols used by mobile nodes. On the other hand, the network layer handover control issues are quite dependent on the network architecture. Accordingly, the mobility issues of FN can be addressed, in particular, in the network-layer perspective.

Annex A
(informative)

Existing IP-based mobility control protocols

A.1 Network-layer mobility control protocols

A.1.1 Mobile IP

Mobile IP (MIP) is a protocol to support the IP mobility and is specified in the IETF. MIP may be divided into Mobile IPv4 (MIPv4) and Mobile IPv6 (MIPv6) per the associated IP version. These two protocols basically provide similar functionality. Details of MIPv4 and MIPv6 are described in IETF RFC 3344 and RFC 3775, respectively.

MIPv4 operates between the following entities: Mobile Node (MN), Home Agent (HA), Foreign Agent (FA) and Correspondent Node (CN). When a MN moves into a new subnet, it registers with the HA with a Care-of Address (CoA). The CoA represents an IP address of FA. The MN must register its CoA with the HA whenever the MN changes its subnet. If the HA receives packets destined for MN from the CN, and the MN is roaming in a visited network, the HA intercepts these packets and forwards them to the CoA through the mobile IP tunnel. The FA decapsulates the received packets from the HA and delivers the original packets to the MN. MIP does not support fast handover for time-critical and loss-sensitive applications. To address this problem, the MIP has been extended to Fast Handover for MIP (FMIP) and Hierarchical MIP (HMIP).

In HMIP, the access networks are organized hierarchically. Gateway Foreign Agents (GFAs) of MIPv4, or Mobility Anchor Points (MAPs) in MIPv6, are responsible for mobility management of mobile nodes within the local domain. Therefore, the movement of mobile modes within the local domain will be hidden from the HA and CN in the other networks, and thus the registration latency and signalling overhead can be decreased considerably. The HMIP architecture for MIPv4 is also known as 'regional registration'.

In FMIP, the MIP registration procedures can start only after the link layer handover is complete. It is noted that, if appropriate information could be obtained from the lower-layer (before the link layer handover is completed), the MIP handover latency could be reduced. This is the main concept of the FMIP approach. In addition, a bidirectional tunnel between access routers can be used to support low loss and low latency handover.

A.1.2 Proxy Mobile IP

MIPv6 requires client functionality in the IPv6 stack of a MN. Exchange of signalling messages between MN and HA enables the creation and maintenance of a binding between the MN's home address and its care-of address. MIP requires the IP host to send IP mobility management signalling messages to the HA, which is located in the network.

PMIP is a network-based mobility, which is another approach to solving the IP mobility challenge. It is possible to support mobility for IPv6 nodes without host involvement by extending MIPv6 signalling messages between a network node and a home agent. This approach to supporting mobility does not require the MN to be involved in the exchange of signalling messages between itself and the HA. A proxy mobility agent in the network performs the signalling with the home agent and does the mobility management on behalf of the MN attached to the network.

Network deployments that are designed to support mobility would be agnostic to the capability in the IPv6 stack of the nodes. IP mobility for nodes that have mobile IP client functionality in the IPv6 stack as well as those nodes that do not would be supported by enabling PMIPv6 protocol functionality in the network.

A.1.3 Hierarchical Mobile IP

To support mobility to mobile devices that are contently moving, the mobile device will need to send signalling messages to the network to maintain reachability. If the mobile device is further away from the home network, many problems will occur such as increase number of messages in the core network, increase of message exchange between the external networks, relatively long propagation delay for handover processing, and disruption to the active connection during handover which may result in data loss. As the number of mobile devices increase, the network will be overflown with control messages rather than user data itself. Thus, it is important to decrease the number of mobility signalling and reduce the delays during the time-critical handover period.

The Hierarchical Mobile IP (HMIP) can be used to reduce the mobility signalling overhead, which was proposed in the IETF RFC 5380. In HMIP, a new network agent, called the Mobility Anchor Point (MAP) is used as a local home agent to the mobile node in the localized network.

Figure A.1 shows an overview of Hierarchical Mobile IPv6. The MAP acts as the anchor point of a mobile node to Home Agent and Correspondent Node. When the mobile node is attached to a new domain, (e. g., domain C in Figure A.1), the mobile node will register its location with the MAP, and then MAP will update the location of mobile node to the Home Agent and Correspondent Node. However, when the mobile node moves around within the same domain, the MAP does not need to further update the location of mobile node to the Home Agent and Correspondent Node. By this, the signalling overhead associated with the location update between MAP and Home Agent or Correspondent Node can be reduced.

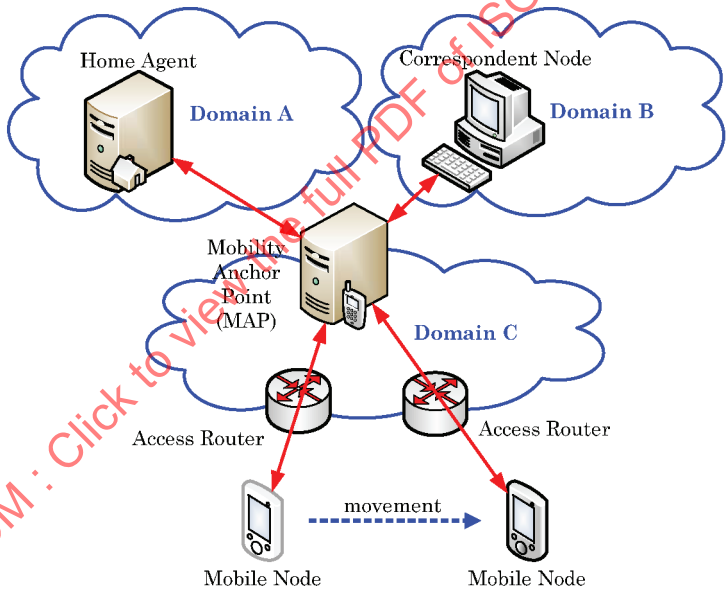

Figure A.1 — Overview of Hierarchical Mobile IP

A.2 SCTP for transport-layer handover support

A.2.1 Overview

The Stream Control Transmission Protocol (SCTP), as defined in the IETF RFC 4960, is the third transport layer protocol next to TCP and UDP. SCTP is featured multi-streaming and multi-homing, differently from TCP. It is noted that the multi-homing feature of SCTP enables the SCTP to support the IP mobility in the transport layer.

More specifically, the SCTP with the dynamic 'Address Configuration (ASCONF)' extension, which is called 'mobile SCTP (mSCTP)', can be used to provide seamless handover for mobile hosts that are moving into different IP network regions during the active session. The mSCTP may be used as an alternative scheme against the handover schemes based on Mobile IP. Differently from the Mobile IP-

based handover schemes, which rely on the support of network routers for tunnelling between access routers, the mobile SCTP provides the handover control at the transport layer without help of routers.

The mSCTP can be used to provide seamless handover for mobile hosts that are moving in to different IP networks. In other words, the mSCTP is targeted for the client-server services, in which the mobile client initiates an SCTP session with the fixed server. For supporting the peer-to-peer services, in which a session is terminated at the mobile host, the mSCTP must be used along with an additional location management scheme such as MIP.

A.2.2 SCTP handover

The SCTP intrinsically provides the multi-homing feature, in which a mobile node is allowed to simultaneously bind multiple IP addresses to its network interface. The recent works on the SCTP include the ASCONF extension. The ASCONF extension enables the SCTP to add, delete and change the IP addresses during active SCTP association.

The SCTP implementation with the ASCONF extension is called the mobile SCTP (mSCTP). The mSCTP can be used for seamless handover while the mobile node is moving into different IP network regions over the session.

Sessions considered in mobile communications can be classified into the following two types:

a) Session originated from mobile host toward fixed host

b) Session originated from fixed host toward mobile host

The mobile sessions in (a) seem to be a natural extension of the client-server model, in which the mobile host originating the session can be viewed as a client, while the counter end point will function as a server. On the other hand, the case (b) requires the additional location management functionality for the session originator to find the current location of the mobile host and to keep track of the location changes, which has so far been addressed by MIP.

The mSCTP, in the present form, is targeted for seamless handover of mobile session associated with the case (a). To support the session type of the case (b), the mSCTP must be used along with an additional location management scheme such as Mobile IP.

Let us consider a mobile client (MC) that initiates an SCTP association with a fixed server (FS). After initiation of an SCTP association, the MC moves from location A (access router A) to location B (access router B). Then, the handover procedures can be performed as follows.

Step 1: Session initiation by MC

We assume that an MC initiates an SCTP association with an FS. The resulting SCTP association is associated with the IPv6 address 2 for MC and the IP address 1 for FS. Note in this phase that the FS is in the single homing with IP address 1. The MC is also in the single-homing in the initial state, in which the IP address 2 is set to its primary IP address in the SCTP initiation process.

Step 2: Obtaining a new IP address for a new location

Let us assume that MC moves from AR A to AR B and thus it is now in the overlapping region. In this phase, we also need to assume that the MC can obtain an IP address 3 from the AR B.

By SCTP implementations, the newly obtained IP address 3 must be signalled or informed to the SCTP in the transport layer, and then the SCTP will bind the new IP address to its address list managed by the SCTP association.

Step 3: Adding the new IP address to the SCTP association

After obtaining a new IP address, the MC's SCTP informs MC that it will use a new IP address. This is done by sending SCTP Address Configuration (ASCONF) Chunk to the FS. The MC may receive the responding ASCONF-ACK Chunk from the FS.

The MC is now in the dual homing state. The old IP address 2 is still used as the primary address, until the new IP address 3 will be set to be "Primary Address" by the MC. Before the primary address is newly set, IP address 3 will be used as a backup path.

Step 4: Changing the primary IP address

While the MC further continues to move toward AR B, it needs to change the new IP address into its primary IP address according at an appropriate rule. Actually, the configuration of a specific rule to trigger this "primary address change" is a challenging issue of the mSCTP.

If once the primary address is changed, the FS will send the incoming data over the new primary IP address, whereas the backup (old) address may be used to recover the lost data chunks.

Step 5: Deleting the old IPv6 address from the SCTP association

As the MC progresses to move toward AR B, if the old IP address gets inactive, the MC must delete it from the address list. The rule for determining if the IP address is inactive may also be implemented by using additional information from the underlying network or physical layer.

Step 6: Repeating the handover procedures

The procedural steps for seamless handover described above will be repeated whenever the MC moves to a new location, until the SCTP association will be released.

A.3 SIP-based mobility schemes

In the application layer, the Session Initiation Protocol (SIP) can be used to provide the location management and handover support functions.

A.3.1 SIP-based location management

The SIP has been made in the IETF for supporting the control of IP-based multimedia sessions as a signalling protocol. SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions. SIP uses SIP Uniform Resource Identifiers (URIs), which are similar to e-mail addresses, as its addressing scheme. It operates independently of the underlying transport layer protocols such as Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).

SIP provides the location management functionality for mobility support based on the user's registration with a SIP registrar. When a SIP User Agent (UA) moves into a new network region, it registers its current location with the location database via a SIP registrar. The location database is referred to by the SIP proxy server or redirect server during UA originated or terminated session initiation.

The SIP functional entities include the UA, the proxy server, the redirect server, the registrar and the location database. SIP messages are classified into two types: request that is sent from the UA Client (UAC) to the UA Server (UAS), and response that contains the status of the request. More details on SIP are given in IETF RFC 3261.

A.3.2 SIP-based handover

The SIP can also be used to support handover using the SIP Re-INVITE message (see IETF RFC 6141 for more details). In SIP handover, a mobile node (MN) performs IP handover by sending another INVITE (called re-INVITE) message to the correspondent node (CN) after getting a new IP address.

The normal SIP handover can be depicted in Figure A.2.

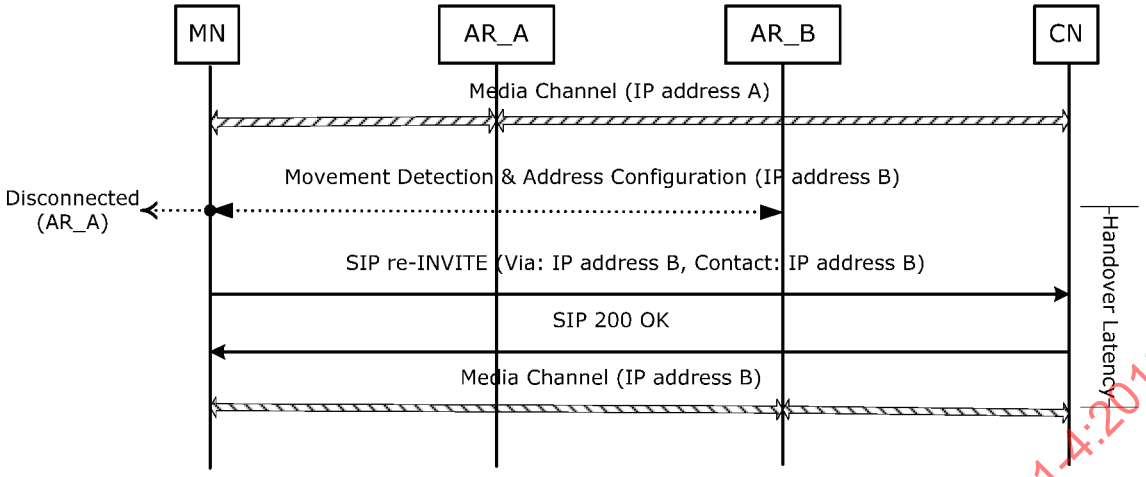

Figure A.2 — Information flow of SIP handover

In the AR_A region, MN is communicating with CN by using IP address A. As it moves into AR_B region, the media channel with IP address A will be disconnected. MN then begins movement detection and address configuration. After getting a new IP address, MN sends an SIP re-INVITE message, which contains the new IP address, and receives the SIP 200 OK message from CN. During this handover period, MN cannot receive the media stream from CN, which induces the concerned handover latency and packet loss.

This normal SIP handover may give a large handover latency associated with movement detection and IP address configuration. Accordingly, the SIP may be extended to support the soft handover with 'bicasting.' In this handover scheme, an MN will communicate with the CN using bicasting in the handover region.

In the SIP handover with bicasting, an MN is assumed to exploit the link-layer triggers such as Link-Up and Link-Down. That is, when the MN goes into the handover region, it will initiate the SIP handover operations with the help of link-layer triggers. On the other hand, it is noted that the existing SIP handover does not use such link-layer information, since the SIP cannot support the soft handover with bicasting. The procedures of SIP handover with bicasting are illustrated in Figure A.3.

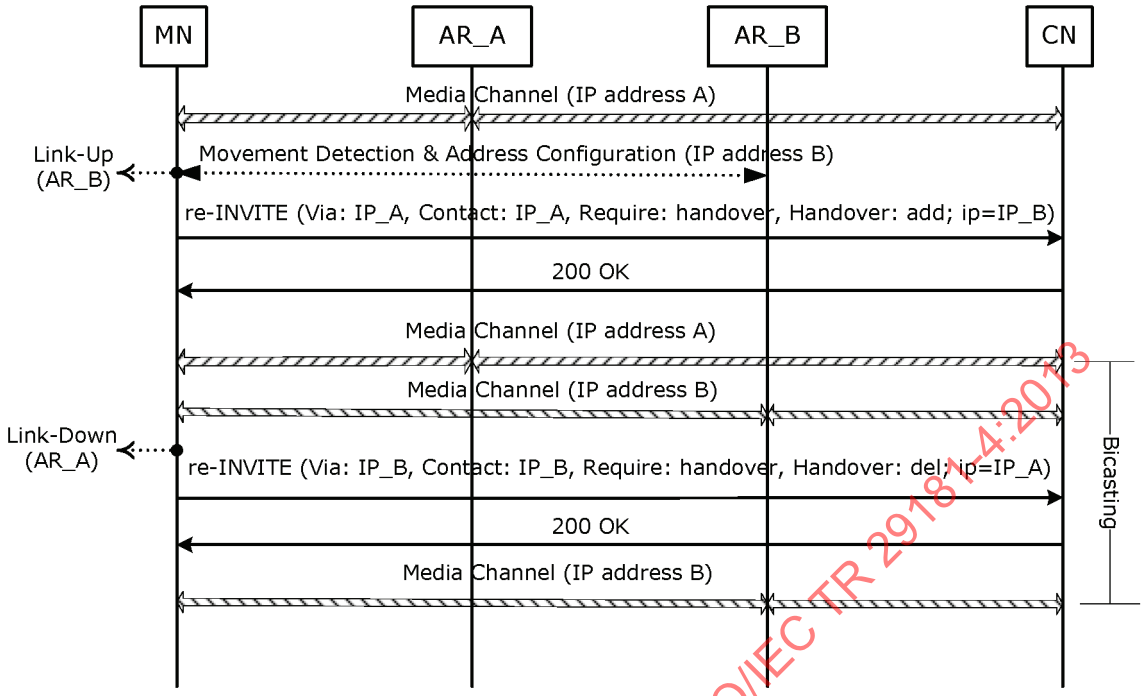

Figure A.3 — Information flow of SIP handover with bicasting

In the figure, MN initially uses IP address A (IP_A). When it moves into AR_B region, it will detect a Link-Up for AR_B. It then performs movement detection and obtains a new IP address (IP_B) via DHCP or IPv6 address auto-configuration.

After getting a new IP address, MN sends an SIP re-INVITE method which contains the information on IP_B and handover header, as specified in the figure. The CN will respond with SIP OK message to MN. Since then, CN can transmit an identical media stream to MN over both IP_A and IP_B. That is, CN starts bicasting to MN. In this period, MN transmits its own media stream to CN using either IP_A or IP_B.

As the MN further moves into the AR_B region, it will detect the Link-Down event for AR_A. The MN then sends an SIP re-INVITE message to CN, as specified in the figure, so as to stop bicasting (i.e. transmission over IP address A). After the corresponding OK message is received, MN and CN use only the IP_B address. The implementation issues of SIP with bicasting are for further study.

Annex B
(informative)

High-level architecture of mobility control in FN

This Annex gives an example of high-level architectures of mobility control in FN, which may be considered in the design of overall FN architecture in the mobility perspective.

B.1 Overview

The mobility control architecture for FN may be designed with three main functional blocks, as shown in Figure B.1, which include ID-LOC separation with Global Identifier and Local Locator (GILL); Global ID-based communication and Local LOC-based delivery (GIC-LLD) in the data plane; Distributed Mobility Control (DMC) in the control plane.

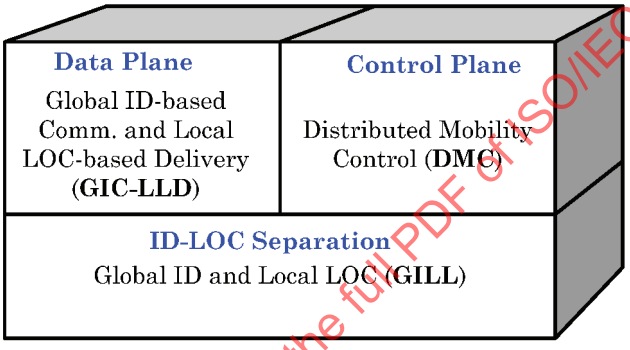

Figure B.1 — Three functional blocks for mobility control

B.2 Global identifier and local locator

In FN, ID is separated from LOC. ID is allocated to a node itself, not its interface. In addition, a LOC is given to the network (rather than a host) that a host is attached to. In particular, LOC is 'local', rather than 'global.' A specific format of ID is for further study. In FN, the mappings between IDs and LOCs will be updated and managed by DMC. In addition, an ID is a global ID that shall be unique in the global network, whereas a LOC is a local IP address that has only to be unique in the local network.

In FN, the IPv4/IPv6 addresses of the access router (AR) and the gateway (GW) may be used as LOCs. These IP address may be private in the network. LOC is used for data delivery in the network, whereas ID is used for end-to-end communication between the two hosts through one or more networks. In mobile environments, a host with a single ID may change its LOCs by movement.

B.3 Global ID-based communication and local LOC-based delivery

In FN, each host has a globally unique ID, by which global communication is accomplished. In the meantime, one or more LOCs are used for packet routing in the network. Each LOC may be used locally in the transit networks, without any assumption on global uniqueness of LOC. Access LOC (A-LOC) is used for forwarding of data packet between hosts and AR in the access network. The format of A-LOC is specific to the underlying access network.

Figure B.2 shows the packet delivery operations with Global ID-based Communication and Local LOC-based Routing in FN.

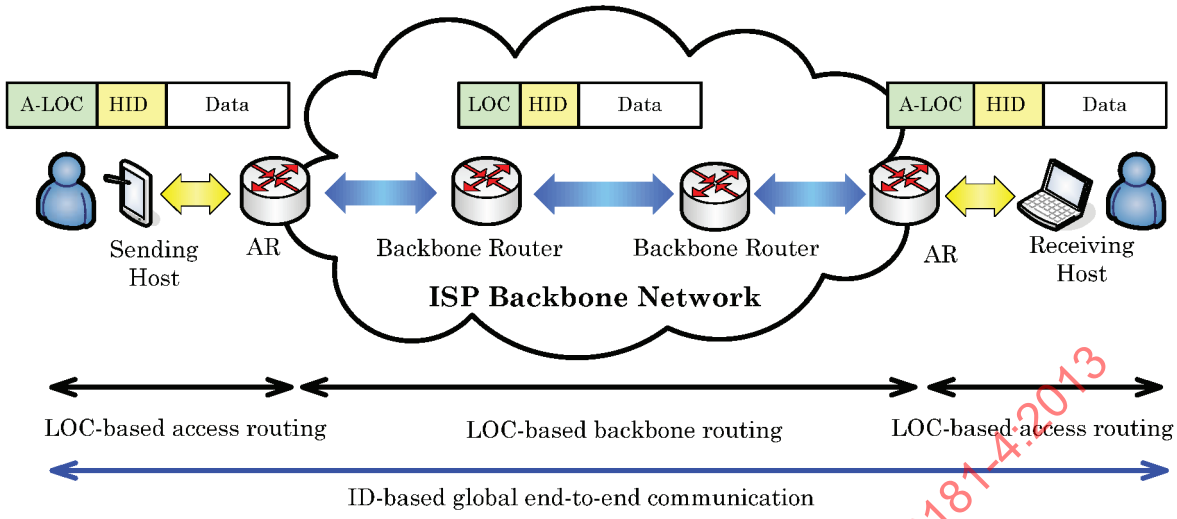

Figure B.2 — Global ID-based communication and local LOC-based delivery

For discussion, we consider a simple data delivery scenario in which a sending host (SH) wants to communicate to a receiving host (RH). It is assumed that SH has the information on ID of RH. Then, the overall data delivery operations a domain are performed as follows.

1) Data transmission (SH ⇔ AR)

   SH sends the data packets to AR in the access network by using ID of RH. It is noted that SH does not need to know the LOC of RH (IP address of AR that is attached to RH). The data packet contains the ID for communication, and it will be delivered to AR, which is based on the A-LOCs of SH and AR.

2) LOC query and LOC translation (AR of SH)

   On reception of data packets from SH, the AR of SH will first investigate the ID of RH, and then identify the LOC of RH by using the DMC. Then, AR of SH will translate the LOC of access network to that of backbone network.

3) Packet delivery in the backbone network

   The encapsulated data packets are delivered from AR of SH to AR of RH, possibly via one or more routers in the backbone network.

4) LOC translation (AR of RH)

   On reception of the encapsulated data packets, the AR of RH will extract the original data packets, and then forward them to RH over the access network. Note that the access network of RH may be different from the access network of SH.

5) Data reception (AR ⇔ RH)

   Finally, the RH can receive the original data packets transmitted by SH.

In FN, data communication will be accomplished based on ID, not IP address that is used in the current Internet. More specifically, a host will initiate a communication session with a ID of the corresponding host. On the other hand, data routing is performed locally in the access and backbone network.

B.4 Distributed mobility control

Most of the existing IP-based mobility protocols are based on the centralized approach, as shown in the MIP and PMIP, in which all control and data traffic will be processed by a centralized mobility anchor, such as Home Agent (HA) of MIP or Local Mobility Anchor (LMA) of PMIP. However, such a centralized

mobility scheme is vulnerable to some problems. First, the centralized anchor may induce unwanted traffic into the core network, which tends to give a big burden to mobile network operators in terms of operational costs. In addition, a single point of failure of the central anchor may affect overall data transmission and degradation of performance, which will increase the cost of network dimensioning and engineering.

To overcome the limitations of centralized mobility control, we may consider the distributed mobility control in FN. In the centralized mobility control, the routing path through a centralized anchor tends to be longer, which results in non-optimal routes and performance degradation, whereas the route optimization will be intrinsically supported in the distributed mobility control. Moreover, the distributed mobility control can reduce unnecessary traffics, if the two end hosts communicate directly each other, not relying on a centralized anchor. This will also be helpful to reduce the handover delay. Moreover, the centralized approach is vulnerable to a single point of failure, whereas the distributed approach will mitigate such problem to a local network.

For distributed mobility control, we divide the network reference model into intra-domain and inter-domain cases. Herein, a domain represents the network domain that is administrated by an ISP.

Figure B.3 shows the network model for intra-domain mobility control, in which data and control flows are illustrated within an ISP domain.

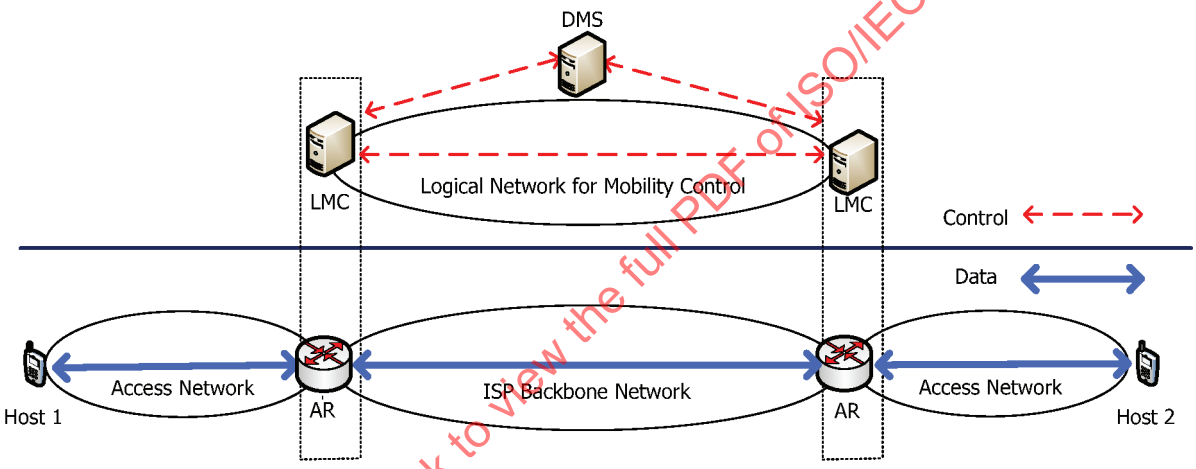

Figure B.3 — Network model for intra-domain mobility control

In the figure each host is attached to Access Router (AR), and a lot of ARs are interconnected with each other in the ISP backbone network. It is assumed that each AR is connected to its own Local Mobility Controller (LMC) via an internal interface. LMC may be implemented over AR. Each domain has a Distributed Map Server (DMS) to perform the mobility control operations. DMS may be implemented over the gateway (GW) of the domain.

LMC is an agent located with AR, which is in charge of control operations for mobility control in the local domain. Each LMC is likely to be implemented with AR. That is, AR and LMC may be just logically (or functionally) separated, but physically co-located over the same equipment. For inter-domain mobility control, each LMC performs the mobility control operations with DMS.

DMS is an agent located with GW, which is in charge of control operations for mobility control in the global domain. For inter-domain mobility control, each DMS performs the mobility control operations with its LMCs and the other DMSs.

Figure B.4 shows the network model for inter-domain mobility control, in which the hosts may be located across different ISP domains. It is assumed that each ISP is interconnected with the other ISPs over global Internet domain.