Mobile SCTP for IP Mobility Support in All-IP Networks

Seok Joo Koh

sjkoh@cs.knu.ac.kr

Abstract

The Stream Control Transmission Protocol (SCTP) is a new transport protocol that is featured multi-streaming and multi-homing. Until now, the two approaches have been considered for IP mobility management: Mobile IP in the network layer and SIP in the application layer scheme. In this paper, we discuss the use of SCTP for IP mobility support in the transport layer. We provide a framework of the use of SCTP with an ADDIP extension, which is called 'mobile SCTP (mSCTP)' in this paper, for seamless handover and location management. Some implementation and deployment issues are also discussed in all-IP wireless mobile networks.

1. Introduction

At the time of the fixed-mobile convergence toward all-IP networks, the IP mobility management issues have been focused in the next-generation wireless mobile networks that include WLAN and Cellular systems [1 - 3].

The mobility management issues can be divided into Location Management and Handover Management. Location Management is used to identify the current location of a mobile node and to keep track of its changes as it moves on. Basically, the location management is done so as to prepare the call setup for sessions that are requested to mobile nodes (MN) by correspondent nodes (CN). With help of location management, the CN will be able to locate the current point of attachment of MN and then establish a session via an appropriate call setup process. On the other hand, Handover Management is used to provide seamless handover for mobile hosts that are moving into different IP network regions during a session. The main objective of the seamless handover is to minimize the service disruption due to data loss and latency during the handover period.

Until now, for IP mobility support, the two approaches have been considered: Mobile IP as a network layer scheme [4 - 8] and SIP as an application layer scheme [9]. Mobile IP [4 - 5] and its enhancements [6 - 8] provide Location Management and Handover Management as well, for 'terminal' or 'host' mobility support in the network layer, whereas Session Initiation Protocol (SIP) is targeted only for providing Location Management for 'user' or 'personal' mobility services in the application layer [9].

In Mobile IPv4 [4], Home Agent (HA) and Foreign Agent (FA) are employed for location management. The Home Address (HoA) is used as a host identifier of an MN, whereas the

Care-of Address (CoA) is used as a location identifier of the MN. The HA maintains the information on the current location of each mobile node by binding CoA of FA to HoA of MN. When a Collaborated CoA (CCoA) is used instead of CoA, the FA is not employed in MIPv4.

In Mobile IPv6 [5], similarly to MIPv4, HoA and CoA are used as a host identifier and a location identifier of the MN, respectively. The CoA of MIPv6 is the same as CCoA of MIPv4. By using MIPv6, an external CN will be able to establish a session with an MN. In Mobile IP, it is noted that the HoA is used, as well as a host identifier, by the applications of an MN, since the MN binds its applications to the HoA. That is, HoA is also used for IP packet data transport.

For handover management for seamless handover, some extensions to Mobile IP are being developed such as the Low Latency handover for MIPv4 [6] and Fast Handover for MIPv6 [7]. These MIP-based handover schemes rely on the tunneling between old and new Access Routers (ARs).

The SIP provides the location management for personal mobility by using a suite of SIP servers such as Registrar, Proxy and Location servers. In response to a session initiation request, the SIP servers will locate the corresponding SIP peer and establish an SIP session. The handover between different network regions during a session is outside the scope of SIP.

The Stream Control Transmission Protocol (SCTP) [10 - 12] is a new transport protocol located at the side of TCP and UDP. Differently from TCP and UDP, the SCTP is featured 'multi-streaming' and 'multi-homing'. In particular, the multi-homing feature of SCTP enables the SCTP to be used for IP mobility support [13 - 17].

In this paper, we propose a framework of the use of SCTP for IP mobility support. In particular, we discuss the mobile SCTP (mSCTP), which is defined as the SCTP with an ADDIP extension [13]. The mSCTP can be used to provide seamless handover for the mobile sessions that are originated by Correspondent Nodes (CN) toward Mobile Nodes (MN) during the session, without support of routers in the networks [14, 15]. We also describe that the mSCTP can be used for the mobile sessions that are initiated by CN toward MN, if it is used along with the Mobile IP or SIP for location management [16]. In this paper, we will first focus on the use of Mobile IP for location management. The mSCTP with SIP is for further study. In case that Mobile IP is used for location management, the CN will find the current location of MN with help of the MIP Home Agent and then establish an SCTP association.

This paper is organized as follows. Section 2 describes an overview of mobile SCTP. In Section 3, we present the procedures of the mSCTP handover for a mobile session that originated by MN to CN, along with some implementation issues. Section 4 discusses the use of mSCTP with Mobile IP for location management, which is targeted for the mobile sessions that are initiated by CN toward MN. In Section 5, the mobile SCTP is compared with Mobile IP and SIP in terms of IP mobility support. Section 6 concludes this paper with future works.

2. Mobile SCTP

Stream control transmission protocol (SCTP) is an end-to-end, connection-oriented protocol that transports data in independent sequenced streams. SCTP endpoints support multi-homing; therefore, interface redundancy is built into the protocol. Through selective transmission mechanisms, SCTP resolves errors and buffers the data transmission process [17].
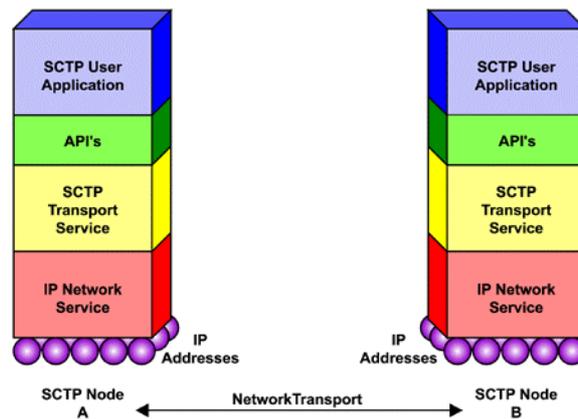

Figure 1 – SCTP Overview

As illustrated in Figure 1, the SCTP transport service layer is positioned between the SCTP user application and the network service being used. Since SCTP is based on interfacing two SCTP endpoints, there are certain application programming interfaces (APIs) that run in between the transport service layer and SCTP user layer. In addition, each endpoint hosts multiple IP addresses. Within SCTP, user data and control information are assembled in chunks.

The multi-homing feature enables SCTP endpoints to support multiple IP addresses. Multi-homing protects an association from potential network failures by steering traffic to alternate IP addresses. During the initiation of an association, SCTP endpoints exchange lists of IP addresses. Therefore, each endpoint can send and receive messages from any of the IP addresses listed at the remote endpoint. For example, one of the listed IP addresses will be designated as the primary address during the initiation. If the primary address repeatedly drops chunks, however, all chunks will be transmitted to an alternate address.

The recent works on SCTP include an ADDIP extension for multi-homing support during a association [13]. The ADDIP extension enables the SCTP to add a new IP address, to delete an unnecessary IP address and to change the primary IP address used for the association, while an SCTP association goes on.

In this paper, we define "mobile SCTP (mSCTP)" as an SCTP implementation with its ADDIP extension. The mSCTP can be used for seamless handover of sessions that are triggered by mobile nodes. It can also be used for mobile sessions that are initiated by fixed nodes toward mobile nodes, together with a location management scheme such as Mobile IP.

3. mSCTP for Seamless Handover

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

    i.   Session originated by mobile host toward fixed host

    ii.   Session originated by fixed host toward mobile host

The mobile sessions in (i) 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 endpoint will function as a server. On the other hand, the case (ii) requires an additional location management to ensure that the session originator is able to find the current location of the mobile host and to keep track of the location changes.

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

3.1 Handover procedures based on mSCTP

In this section, we consider a mobile node (MN) that initiates an SCTP association with a correspondent node (CN). After initiation of an SCTP association, the MN moves from Location A (Access Router A) to Location B (Access Router B), as shown in Figure 2. The figure illustrates an example of the use of mobile SCTP for seamless handover in IPv6 networks. With this figure, we describe the detailed handover procedures in the subsequent sections.
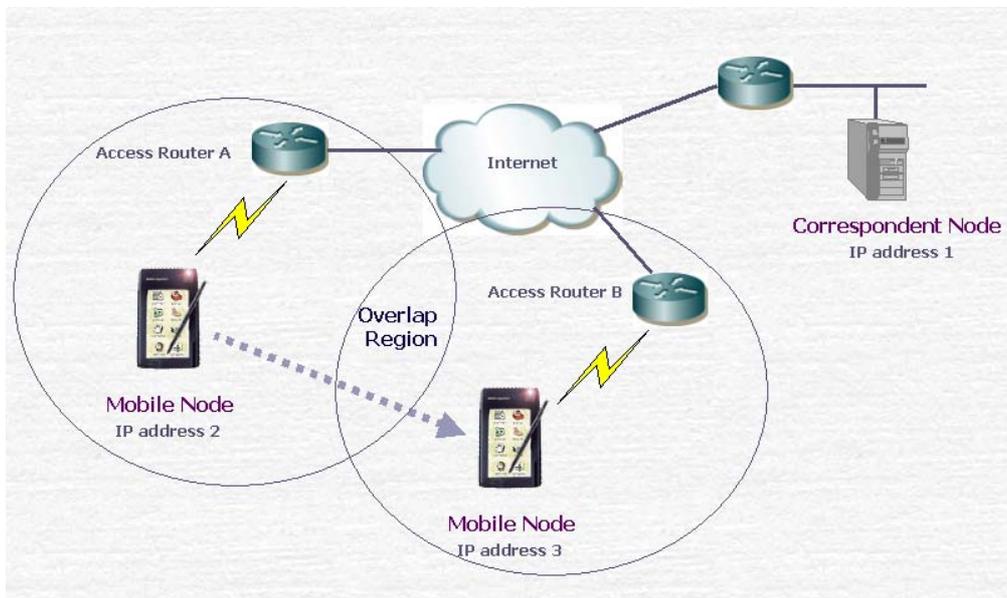

Figure 2 – Mobile SCTP for Seamless Handover

3.1.1 Session initiation by mobile node

We assume that an MN initiates an SCTP association with a CN. The resulting SCTP association has the set of IP addresses with IP address 2 for MN and IP address 1 for CN. It is also assumed that the MN gets an IP address from Access Router (AR) A, with help of IPv6 stateless address auto-configuration or DHCPv6.

Note in this phase that the CN is in the single homing with IP address 1. The MN is also single homing in the initial state, in which the IP address 2 is set to its primary IP address in the SCTP initiation process.

3.1.2 Obtaining an IP address for a new location

Let us assume that the MN 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 MN can obtain a new IP address 3 from the AR B by using DHCPv6 or IPv6 stateless address auto-configuration.

By SCTP implementations, the newly obtained IP address 3 MUST be signaled 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.

3.1.3 Adding the new IP address to the SCTP association

After obtaining a new IP address, the MN's SCTP informs CN that it will use a new IP address. This is done by sending SCTP Address Configuration Change (ASCONF) Chunk [13] to the CN. The MN may receive the responding ASCONF-ACK Chunk from the CN.

The MN is now in the dual homing state. The old IP address (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 MN. Before the primary address is newly set, IP address 3 will be used as a backup path.

3.1.4 Changing the primary IP address

While the MN 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.

Examples of the rules for triggering the primary IP address change are described below:

(a) As soon as a new IP address is detected

When a new IP address is detected, the MN may trigger the primary IP address change by sending the ASCONF Chunk containing the "Set Primary Address" parameter.

This choice is the simplest way to implement in SCTP. In fact, it is the only scheme to apply for Wireless LAN, since in WLAN the multi-homing is not support in the link layer (i.e., Access Points allow only one wireless channel at a time).

This scheme is preferred in terms of the handover latency, in particular, for the fast-moving MNs. However, this is not suitable in cases that MNs shows a so-called ping-pong effect between two neighboring locations.

(b) By using a threshold for the data loss experienced

This rule can be applied when the SCTP of MN has a pre-configured threshold for data loss. For example, if the number of the lost data chunks is greater than the pre-configured threshold, then it may trigger the primary address change toward the CN. The selection of an appropriate threshold value is for further study, and may depend on implementations and the mobility behavior considered.

This scheme is preferred in the viewpoint that it can operate without support of the underlying network and link layers. However, it is required that SCTP implementation should be extended so as to handle the threshold scheme.

(c) By using an explicit indication from the underlying layer (by comparison of strength of the detected wireless signals)

If the underlying wireless link layer can detect and compare the signal strength of the physical media, and if it can also inform the SCTP about a certain indication (i.e., by using a up-call), then the MN may trigger the primary address change according to such an indication.

This rule is a more preferred choice, if possible, in terms of handover efficiency. It however depends on the capability of the underlying wireless systems.

If once the primary address is changed, the CN 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.

3.1.5   Deleting the old IP address from the SCTP association

As the MN progresses to move toward AR B, if the old IP address gets inactive, the MN 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, as described 3.1.4.

3.1.6   Repeating the handover procedures

The procedural steps for seamless handover described above, from 3.1.1 through 3.1.5, will be repeated whenever the MN moves to a new location, until the SCTP association will be released.

3.2 Implementation issues

The only requirement for providing the seamless handover by using SCTP is that the MN and CN should be equipped with mSCTP implementations, (i.e., SCTP with ADDIP extension.)

On the other hand, the following issues need to be considered for implementations.

(a) Number of IP addresses used by CN

In this document, we assume that the CN is in the single homing. As an alternative, especially if the CN is a dedicated server, the CN may strategically assign two or more IP addresses to its network interfaces, so as to enable easy distinction of the two links at the MN. This allows SCTP to represent different links by different entries in the host routing table of the MN. For more information on this issue, please refer to [12, 15].

(b) Dynamic IP address configuration

The basic assumption for seamless handover is that an MN is able to obtain a new IP address from a new location. Typically, this will be implemented by using DHCP in IPv4 networks and DHCPv6 or stateless address auto-configuration in IPv6 networks. We need to examine the handover latency incurred for obtaining the new IP address via DHCP or IPv6 through experiments.

(c) AAA Functionality

It is envisioned that the deployment of mSCTP will be done along with the appropriate AAA server in the respective access network domains. The AAA server is used to authenticate the MN and also to authorize the new IP address configuration for each location. However, this issue is outside the scope of mSCTP.

(d) Multi-homing Support at the wires link layer

It is noted that the multi-homing capability for mobile node actually depends on the characteristics of the concerned wireless links such as WLAN or Cellular. In the systems such as WLAN, it may not be allowed for an MN to simultaneously bind two different Access Points (APs), and thus the multi-homing is not easy to support. On the other hand, the wireless links used in Cellular systems are expected to easily support the link-level multi-homing features.

(e) Measuring the Wireless Signal Strength at the Physical Layer

Part of the mSCTP seamless handover operations depend on the support of the underlying physical and link layers so as to measure the wireless signal strength. The information on measured signal strength can be helpfully used for SCTP to trigger the addition/deletion of IP addresses and the change of the primary address.

4. mSCTP with Mobile IP for Location Management

To support the mobile sessions that are initiated by a CN toward an MN, the mSCTP may be used along with a location management scheme such as Mobile IP (MIP). In this scenario, the MIP will be used for a CN to locate an MN and to establish an SCTP association with the MN. After an SCTP association is successfully setup, the mobile SCTP will be used for providing seamless handover for the MN, as described in Section 3.

If once the association is established, the data transport between MN and CN relies only on SCTP, not using MIP. The tunneling between HA and MN is not used. Furthermore, the Home Address (HoA) of MN is not used for the data transport. Note that the HoA is used for only the location management

4.1 Initiation of an SCTP association

As described in RFC 2960 [10], the SCTP initialization is done between CN and MN as follow:

(a) CN sends INIT chunk to MN for triggering the association setup. The INIT chunk contains a list of IP addresses and port number that will be used by CN.

(b) MN responds with INIT-ACK chunk to CN. INIT-ACK chunk contains a list of IP addresses and port number that will be used by MN. It can also contain indication about which is the primary address among multiple addresses, if any.

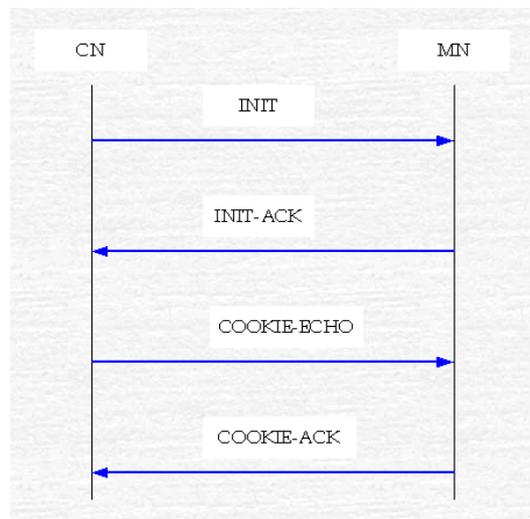(c) CN and MN then exchange COOKIE-ECHO and COOKIE-ACK chunks each other for security enhancements.

Figure 3 – SCTP Initiation Procedures

As shown in Figure 3, the establishment of an SCTP association is triggered by exchanging INIT and INIT-ACK and completed by exchanging COOKIE-ECHO and COOKIE-ACK chunks between CN and MN.

In Mobile IP, HA has information on the current location of an MN, which will be updated by the MIP Registration or Binding Update procedures. The location management of Mobile IP can be used to convey the INIT chunk message from CN to MN via HA.

After receiving the INIT chunk, the MN responds with INIT-ACK chunk directly to CN, not by way of HA. The responding INIT-ACK chunk must contain the CCoA (in MIPv4) or CoA (in MIPv6) that is addressable to the MN.

4.2 mSCTP over Mobile IPv4

Mobile IPv4 considers two options to represent the location of the foreign link: CoA and CCoA. The CoA is the IP address of FA (possibly a router), whereas the CCoA is dynamically assigned to MN by using an address allocation mechanism such as DHCP.

Unfortunately, the CoA of MIPv4 cannot be applicable to SCTP, since it is an address of FA and thus cannot be used within the MN host (for its SCTP association). Note that SCTP is an end-to-end transport layer protocol, not a network-layer one. On the other hand, the CCoA in MIPv4 can be used within the SCTP hosts. Specifically, the SCTP of MN could bind the CCoA to an SCTP association. Therefore, the current usage of SCTP over MIPv4 is restricted to the case in which the MN hosts employ CCoAs in a foreign link. The case of using CoA is for further study.

The detailed procedures for SCTP initiation with MIPv4 are similar to those for the case with MIPv6, which will be described in the next section.

4.3 mSCTP over Mobile IPv6

In MIPv6, the CoA is used instead of CCoA of MIPv4. The CoA may be obtained from the foreign location via DHCPv6 or stateless address auto-configuration. Let us consider an example of MIPv6 networks, which consists of CN, MN and HA. The MN is now at a foreign link, as shown in Figure 4.
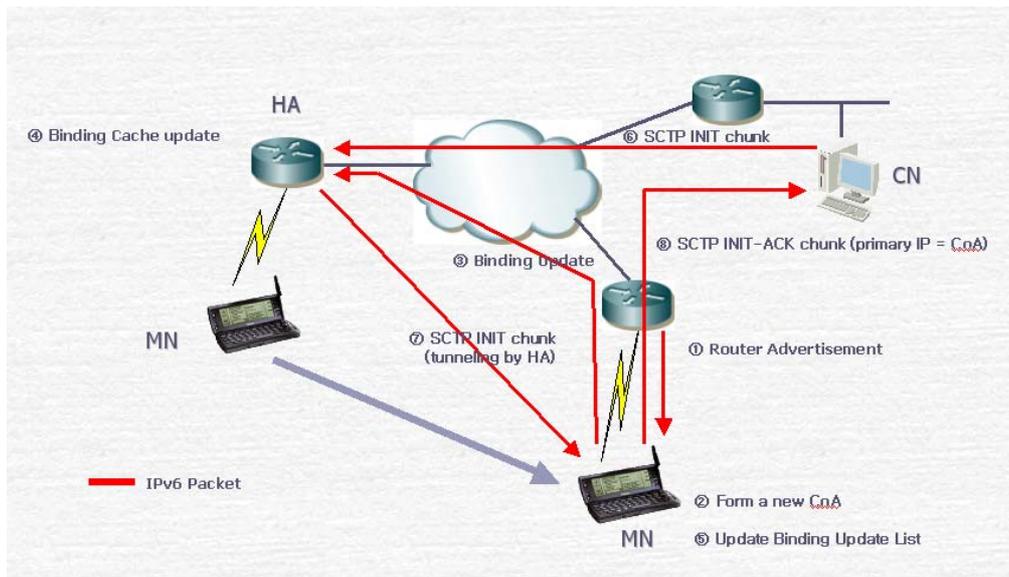

Figure 4 – SCTP Initiation over MIPv6

4.3.1 SCTP initiation over MIPv6

In the example, the MN has already obtained a CoA from the foreign link (as Steps 1 and 2 in the figure), and it also registered its CoA with HA by using the MIPv6 Home Registration Procedures (as done in Steps 3, 4 and 5). These procedures are exactly the same as the MIPv6 registration mechanisms.

Here, we assume that the applications of MN can initially bind the CoA as well as HoA via the socket interface. After initialization of SCTP association, the HoA may be released from the application, as described below. It is noted that the HoA is still used for MN to update its new CoAs to HA according to the MIPv6 mechanisms.

Now, the CN initiates an SCTP association establishment with the MN by sending INIT chunk message over the home address (HoA). The INIT chunk will be first routed to HA, as Step 6 in the figure. In turn, the HA will forward the INIT chunk to MN by referring to CoA and using a tunneling mechanism, as Step 7 in the figure.

In response to the INIT chunk, the MN sends INIT-ACK chunk to the CN, as illustrated in Step 8 of the figure. The INIT-ACK contains the CoA address (as the Primary address) as well as HoA address. Here, the HoA address may only be used for the CN to check whether the responding MN is the authorized host or not (for somewhat security reason). In fact, the HoA will not be referred to by CN (see the Section 3 for more details). Note that the source address and destination address of the IP packet containing the INIT-ACK chunk are CoA of MN and IP address of CN, respectively.

In turn, the COOKIE-ECHO and COOKIE-ACK chunks will be exchanged between CN and MN, so as to complete the procedures of the SCTP association establishment. The SCTP data transport will be done directly between MN and CN without support of network routers.

4.3.2 Seamless handover during association

After the association is established, the CN transmits data chunks to MN over the CoA address, and the MN sends data chunks to CN directly. When the MN changes its points of attachment to the IP networks, the MN will get a new CoA address from the new Access Router. According to the MIPv6 mechanism, the MN will update its new location to HA, which is done regardless of the on-going SCTP association.

On the other hand, the MN will perform the seamless handover, as it moves into a new IP subnet area, according to the mobile SCTP by adding the new IP address to the on-going association, as described in Section 3. These procedures will be repeated until the association has been completed.

5. Comparison of Mobile SCTP, Mobile IP and SIP

In this section we briefly compare the existing mobility schemes with the proposed mobile SCTP, as summarized in Table 1.

Table 1. Comparison of mSCTP, MIP and SIP in terms of IP mobility support

Category	Mobile SCTP	Mobile IP	SIP
Target Layer	Transport	Network	Application
Concerned Transport Services	SCTP	TCP/UDP	TCP/UDP
Location Management Support	No	Yes	Yes
Handover Management Support	Yes (for location management, it can be used with MIP or SIP)	Yes (with MIP extensions such as FMIPv6)	No (for handover, it can be used with mSCTP)
Route Optimization	Intrinsically provided	Provided by binding update with CN	No
Router Support	Not necessary	Necessary	Not necessary
Deployment Perspective	SCTP is required for end hosts (MN & CN)	MIP is required for networks routers and hosts	SIP is required for end hosts and SIP servers are needed

As described in Table 1, the mobile SCTP can be used for seamless handover in the transport-layer. To use mSCTP, it is required that the CN and MN hosts should be aware of the mobile SCTP. Instead, mSCTP does not need the support of network routers for seamless handover. Furthermore, the mSCTP intrinsically provides the Route Optimization without using any additional Binding Update procedures.

For location management, the mSCTP may be used along with MIP or SIP. In case of using MIP for location management, only the MN needs to be aware of MIP, whereas the CN need not use MIP. Using mSCTP with MIP, the MN must also be able to bind the CoA as well as HoA to its applications. The HoA will be used only for location management. After establishment of an SCTP session, the HoA will not be used for data transport. Instead, the CoA is employed for the SCTP data transport. On the other hand, in MIP, only HoA is bound to the applications of MN regardless of the different CoAs.

The MIP provides the location and management in the network layer, and it can support seamless handover with help of neighboring routers such as tunneling between old and new ARs. On the other hand, SIP is an application signaling protocol that supports the location management for user or personal mobility. SIP itself does not provide seamless handover. It may be used together with mSCTP for seamless handover.

6. Conclusions and Further Issues

In this paper, we described a new scheme for IP mobility support, mobile SCTP (mSCTP). The mSCTP can be used to provide seamless handover for mobile sessions that are originated from mobile nodes, without any support of network routers. The mSCTP can also be used for the mobile sessions that are originated from a fixed node to a mobile node along with an appropriate location management scheme such as Mobile IP or SIP. In this paper, we describe a framework of the use of SCTP for IP mobility support.

We note in mSCTP that the home address (HoA) of MN is not involved in the data transport between CN and MN. The reason for this is to exploit the intrinsic route optimization feature of mobile SCTP. On the other hand, the MIP-based applications are bound to the HoA in the MN.

For further study, we need to consider the use of HoA in SCTP over MIP. In this case, the HoA may be used as a backup IP address against the event of path failure of the primary CoA address. This scheme will however require the additional tunneling between HA and MN, and binding update between CN and MN.

It is noted that the handover performance of mSCTP seems to severely depend on the support of the wireless link layer. If interaction from L2 to L3 and L4 is defined appropriately, the SCTP handover seems to be more improved. The performance may also be affected by thetime duration that the MN stays in the overlapping region and the capability of the link layer support for multi-homing. From the point of view of current practice, the 3G cellular networks rather than the Wireless LAN seem to be more suitable to mSCTP, since WLAN does not support the link-layer multi-homing.

References

[1] Lin Y-B. and Chlamtac I., Wireless and Mobile Network Architecture, John Wiley & Sons, Inc., 2001

[2] Evans B. G. and McLaughlin S., "Visions of 4G", IEE Electronics & Communication Engineering Journal, December 2000

[3] Ala-Laurila J., et al., "Wireless LAN Access Network Architecture for Mobile Operators", IEEE Communication Magazine, November 2001

[4] Campbell A. T., et al., "Comparison of IP Micro-Mobility Protocols", IEEE Wireless Communications Magazine, February 2002

[5] Perkins, C. (ed.), "IP Mobility Support for IPv4", IETF RFC 3344, August 2002

[6] Johnson, D., et al., "Mobility Support in IPv6", IETF Internet Draft, draft-ietf-mobileip-ipv6-20.txt, January 2003

[7] Malki, K. L., et al., "Low Latency Handoffs in Mobile IPv4", IETF Internet Draft, draft-ietf-mobileip-lowlatency-handoffs-v4-04.txt, June 2002

[8] Koodli, R., et al., "Fast Handovers for Mobile IPv6", IETF Internet Draft, draft-ietf-mobileip-fast-mipv6-05.txt, September 2002

[9] Schulzrinne H. and Wedlund E., "Application-Layer Mobility using SIP", ACM Mobile Computing and Communications Review, Vol. 4, No. 3, July 2000

[10] Stewart R., et al., "Stream Control Transmission Protocol", IETF RFC 2960, October 2000

[11] Ong L. and Yoakum J., "An Introduction to the Stream Control Transmission Protocol (SCTP)", IETF RFC 3286, May 2002

[12] Coene L., "Stream Control Transmission Protocol Applicability Statement", IETF RFC 3257, April 2002

[13] Stewart, R., "Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration", IETF Internet Draft, draft-ietf-tsvwg-addip-sctp-05.txt, May 2002

[14] Riegel, M. and Tuexen M., "Mobile SCTP", IETF Internet Draft, draft-riegel-tuexen-mobile-sctp-02.txt, February 2003

[15] Koh S. J., et al., "Use of SCTP for Seamless Handover", IETF Internet Draft, draft-sjkoh-mobile-sctp-handover-00.txt, February 2003

[16] Koh S. J., et al., "SCTP with Mobile IP for IP Mobility Support", IETF Internet Draft, draft-sjkoh-mobile-sctp-mobileip-00.txt, February 2003

[17] SCTP tutorial, http://www.iec.org/online/tutorials/sctp/