



Georg-August-Universität
Göttingen
Institut für Informatik

ISSN 1611-1044
Nummer IFI-TB-2005-001

Technischer Bericht

A Review of Mobility Support Paradigms for the Internet

Deguang Le, Xiaoming Fu, Dieter Hogrefe

**Technische Berichte
des Instituts für Informatik
an der Georg-August-Universität Göttingen**

January 2005

Georg-August-Universität Göttingen
Institut für Informatik

Lotzestraße 16-18
37083 Göttingen
Germany

Tel. +49 (5 51) 39-1 44 14

Fax +49 (5 51) 39-1 44 15

Email office@informatik.uni-goettingen.de

WWW www.ifi.informatik.uni-goettingen.de

A Review of Mobility Support Paradigms for the Internet

Deguang Le Xiaoming Fu Dieter Hogrefe
Telematics Group, University of Göttingen
E-mail: {le,fu,hogrefe}@cs.uni-goettingen.de

Abstract

With the development of mobile communication and Internet technology, there is a strong need to provide connectivity for roaming devices to communicate to other communication end point in the Internet at any time and anywhere. The key issue of this vision is how to support mobility in TCP/IP networks. In this paper, we review the TCP/IP protocol stack and analyze the problems associated with it in a mobile environment. We then investigate the mobility support techniques and existing solutions to provide mobility support in the Internet. We classify the proposed solutions based on the protocol layers and present examples for each category. We also provide a comparison of the different solutions belonging to different categories and in the same category, including their advantages and disadvantages, and conclude that there is no single solution perfectly addresses mobility support for the Internet. We conclude this survey with a recommendation of features that need to be satisfied in Internet mobility support.

1. Introduction

With the increasing penetration of more and more mobile devices which demand for accessing to the Internet and get information and services at any time and anywhere, there is a need for the Internet infrastructure to provide mobile devices with a capability of connecting to the Internet while roaming, preferably without interruption and degradation of communication quality, which we will study in this paper.

Since the existing Internet was originally designed for communications between fixed nodes, Internet mobility support is a very complicated topic, and there are a lot of issues to be resolved. Since last decade, studies that address these issues have arisen, coming up with a number of protocol proposals and even IETF RFCs. Many of them have been designed, implemented and some of them are starting to be deployed. Nevertheless, as analyzed in more detail below, they have both pros and cons in dealing with mobility support in terms of efficiency, functionality and security. Given the importance of mobility support in the Internet, there have been many new protocols and solutions such as SCTP, DCCP, MAST and MOBIKE etc. developed for Internet mobility. These new protocols and solutions proposed new idea and

technique to benefit mobility support. However, most existing proposals (including the IETF Mobile IP starting to be deployed) still suffer from a number of concerns. Therefore, it necessitates a general comparison of different solutions including new emerged alternatives, and a review and rethinking of the architectural aspect of Internet mobility support become necessary. Among previous works, Henderson [1] presented three host mobility solutions, namely MIP, TCP Migrate, and HIP operating in different layers, and compared them from various aspects of performance, security, deployment, scalability, and robustness properties etc. Eddy [2] discussed the strengths and weaknesses of implementing mobility at three different layers of TCP/IP stack, suggesting that the transport layer is the probably the best layer candidate to accommodate Internet mobility, and that there should be more collaboration between layers to avoid confliction and inefficiency. These exiting works did discuss some of existing and emerging mobility approaches and propose some interesting metrics for comparison, however their reviews mainly focused on high layer overview while an in-depth analysis of the underlying properties of introducing mobility to TCP/IP architecture and different mobility support paradigms was not missing; moreover, some other approaches are not considered. The objective of this paper is then to investigate and compare existing Internet mobility support paradigms as comprehensively as possible.

This paper is organized as follows. In section 2, we review the traditional TCP/IP protocol stack, and present some general goals for any solution to mobility support for the Internet. In particular, we describe characteristics of communications in mobile environments, the performance requirements for Internet mobility support, and why a traditional TCP/IP network is unable to support mobility. Section 3 presents a detailed set of mobility support paradigms, each representing some specific changes to the existing protocol layer, and studies the possible effect and impact, especially the integration of different mobile support paradigms. In Section 4, we summarize the advantages and disadvantages introduced by these different paradigms, and conclude that all existing solutions have different implications to their application scenarios; there is no single perfect solution so far; mobility support may require some rethinking of the Internet architecture, and there should be some

general design considerations for any Internet mobility support solution.

2. The TCP/IP Stack and Why Mobility Support is Difficult

In this section, starting with a review of the traditional TCP/IP stack, we describe the problems, requirements and general goals for introducing mobility support in this stack.

2.1. TCP/IP Stack: a Review

In Internet communications a number of protocols have to be run in both end hosts and routers, utilizing a four-layer architecture (see Fig. 1), where the Transmission Control Protocol (TCP) and the Internet Protocol (IP) are fundamental elements in the architecture (hence it is called the TCP/IP stack).

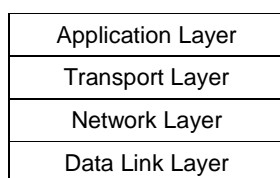


Figure 1: TCP/IP Protocol Stack

In TCP/IP, the application layer defines TCP/IP application protocols and how user programs interface with transport layer to use the network. It provides the services that user applications use to communicate over the Internet. The transport layer provides an end-to-end delivery service: TCP provides a connection-oriented service which allows reliability, fragmentation and flow control and congestion control, while UDP provides an unreliable datagram service that enhances network functions. The network layer is responsible for the routing and delivery of data across networks of the same and different types from a source host towards a destination host. The lowest layer, the data link layer, handles issues concerning the physical addressing, network topology, error notification, sequencing of frames, and flow control between neighboring nodes.

2.2. Basic Functional Requirements for Internet Mobility Support

Here, Internet mobility refers to that an IP-based device (host) moves (i.e., changes its topological point of attachment) to different networks while keeping its active communication(s) continue, excluding the cases where the device just moves within a single network cases (or, data link layer mobility). In order to provide such a support, a number of fundamental problems raise, which can be summarized as following requirements for Internet mobility support.

Handover Management: This requirement is concerned with maintaining the ongoing communication alive while a mobile node (MN) moves and changes its point of attachment. Its main objective is to minimize the service disruption during handover.

Location Management: This involves identifying the current location of the MN as it moves on.

Multihoming: to enhance throughput, is desirable for an MN equipped with multiple interfaces to have multiple communication paths across either the same or different access networks in mobility scenarios.

Applications: Internet mobility should support current services or applications without requiring changes of them.

Security: It is important for an Internet mobility support paradigm to protect itself against misuses of the mobility features and mechanism.

We argue that a complete and useful Internet mobility should address all these requirements. In addition, there are performance requirements for mobile environments as identified as below.

2.3. Performance Requirements for Internet Mobility Support

While developing an Internet mobility solution, the performance metrics also deserve special attention. We consider the following performance metrics that are the most relevant for Internet mobility.

Handover Latency refers to the elapsed time from the last packet received via the old network to the arrival of the first packet along the new network during a handover.

Packet Loss is defined as the number of packets lost for the maintained communication during a handover.

Signaling Overhead is defined as the number of messages for the handover and location procedures.

Throughput is the amount of data transmitted over a mobile Internet in a given period of time.

2.4. Limitation of Traditional TCP/IP for Internet Mobility

The traditional TCP/IP was designed for fixed computer networks and does not well consider the situation of host mobility. This subsection will demonstrate some of the limitations of TCP/IP for Internet mobility.

2.4.1. Limitation of Link Layer. To its maximal possibility, wireless access techniques only provide the mobility of homogeneous networks at the link layer [3], which is not appropriate to Internet mobility across heterogeneous networks. In general the nature of network heterogeneity requires mobility support functions to be provided in higher layers.

2.4.2. Limitation of IP address. The IP address of network layer plays both roles of locator and identifier. In mobile scenarios, the IP address of a roaming device has to be changed to represent the change of its point of attachment to the network when it moves from one to another network. In traditional TCP/IP, a change of the IP address makes it impossible for other devices to contact the device using a constant IP address. In addition, even if the device is able to obtain a new IP address dynamically, the transport connections established in the previous network will break for the change of IP address.

2.4.3. Lack of Signaling Mechanism between Layers. For example, the design of traditional transport layer protocols relies on the services provided by the network layer, and considers by no means the wireless link properties, thus, the congestion control of transport layer does not distinguish the packet loss caused by wireless link properties from the normal packet loss in wired network, which degrades transport performance [4].

2.4.4. Limitation of Applications. Many applications based on traditional TCP/IP architecture are also restrained in their use in mobile environments. For example, in Domain Name System (DNS), a domain name is usually statically bound to a host's IP address, thus the tight binding between domain name and IP address will be invalid because of the dynamic change of IP addresses of mobile devices.

3. Extending TCP/IP for Internet Mobility

As it is mentioned in the previous section, the traditional TCP/IP is not appropriate for Internet mobility. Therefore, various solutions have been developed to address it. Among them Mobile IP [5], [6] and Location Independent Network Architecture for IPv6 (LIN6) [7] represent the network layer ones. In transport layer, a wide range of investigations have been made to provide mobility support for TCP, the Stream Control Transmission Protocol (SCTP) [8], and the Datagram Congestion Control Protocol (DCCP) [9]. Session Initiation Protocol (SIP) [10] Dynamic DNS (DDNS) [11] provide mobility support in application layer.

Some researchers were interested in introducing new protocol layer between classic network layer and transport layer to provide Internet mobility such as Host Identity Protocol (HIP) [12], Multiple Address Service for Transport (MAST) [13], IKEv2 Mobility and Multihoming (MOBIKE) [14].

In this section, we investigate the solutions for improving mobility of TCP/IP in more details.

3.1. Mobility Support in Network Layer

Because IP is the ubiquitous internetworking layer for the Internet, solutions extending the existing network layer are considered a natural approach. MIPv4 proposed by Perkins et al. [5], MIPv6 by Johnson et al. [6] and various enhancements to the performance of MIPv4/v6 proposed in [18-21, 24-26] have formed the "classic" way of supporting mobility in the Internet. LIN6 proposed by Teraoka et al. [7] provides an alternative to mobility support for MIPv6. These protocols take the techniques like proxy, tunnelling [15] and separating [16] etc. to deal with mobility.

3.1.1. Mobile IP and Its Enhancement. Mobile IPv4 (MIPv4) defines a home network where the MN is assigned a permanent IP address called home address which identifies the MN; and it also defines foreign networks where the MN visits. And it introduces two new entities, namely the Home Agent (HA) and the Foreign Agent (FA), to relay the packets between Mobile Node (MN) and Correspondent Node (CN).

In MIPv4, when the MN is on its home network, it acts like any other fixed node of that network and requires no special mobile IP features. Each time when it moves out its home network and accesses to a foreign network, it obtains a Care of Address (CoA) through DHCP [17] etc. and inform its HA of the new address by sending Registration Request Message to the HA. Upon the HA receiving the Registration Request Message, it shall reply to the MN with a Registration Reply Message. The HA then assumes the MN and once a packet destined to the MN arrives at the home network, the HA shall intercept the packet and forward it to the MN with the CoA by tunnelling technique. Once the FA receives the packet it removes it from the tunnel and delivers it to the MN. When the MN wishes to send packets back to the CN, the packets are sent directly from the MN to the destination.

In specifically environment where MN changes their point of attachment to network frequently and the numbers of mobile users grow simultaneously, a number of micro-based mobility protocols such as regional registration [18], Fast Handover for MIPv4 [19], Hawaii [17] and Cellular IP [21] are proposed to improving performance of MIPv4.

In MIPv6, When the MN moves to another network, it acquires the CoA through either stateful [22] or stateless [23] automatic Address Auto-configuration. After obtaining a new CoA the MN registers to HA and CN with a Binding Update messages (BUs), which resolve the triangle routing problem. After this, the flows between MN and CN can be routed directly.

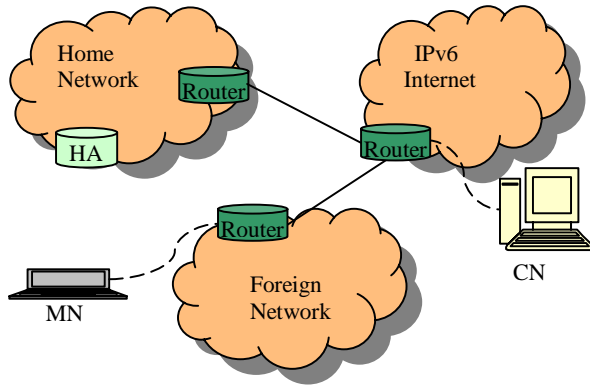


Figure 4: MIPv6 Architecture

Because BUs are transfer between MN and CN as well as HA, this incurs extra overhead. Thus, the IETF develops the Hierarchical Mobile IPv6 (HMIPv6) [24] and Fast Handovers for Mobile IPv6 (FMIPv6) [25] protocols to reduce overload and improve handover speed. Jung et al [26] propose a combination of both approaches of HMIPv6 and FMIPv6, which is designed to add up the advantages of both.

3.1.2. LIN6. LIN6 proposes an alternative Internet mobility solution for IPv6 protocol. Its basic idea is separating identifier and locator in IPv6 address. LIN6 introduces the LIN6 ID for each node as the node identifier so that each node can be identified by its LIN6 ID no matter where the node is connected and no matter how many interfaces the node has. Besides it defines two types of network addresses: the LIN6 generalized ID and LIN6 address. The LIN6 generalized ID is formed by concatenating a constant value called the LIN6 prefix before the LIN6 ID. It is used at the transport layer to identity the connection. The LIN6 address is formed by concatenating the network prefix and LIN6 ID. It is used to routing packet over network layer and the network prefix will change according to the network where the mobile node attaches. Figure 5 illustrates the LIN6 model.

Application Layer
Transport Layer: <LIN6 generalized ID, port> pair
Network Layer: Translation between LIN6 generalized ID and LIN6 IP address
Data Link Layer

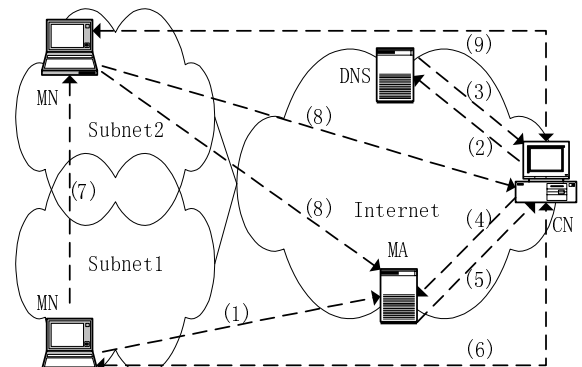
Figure 5: LIN6 Protocol Model

In LIN6, on packet transmission, the network layer extracts the LIN6 ID form the LIN6 generalized ID and concatenates the network prefix and LIN6 ID to create the LIN6 address of the destination node. On packet reception, the network layer remove the network prefix part of the LIN6 address, then attaches the LIN6 prefix

to create the LIN6 generalized ID of the source node. When MN moves to another network and obtain the network prefix of the new network, The MN will update its location with CN in one of two ways: if MN has a security association, it sends the Mapping Update Request message to CN. In this case, the Mapping Update Request message must include the Authentication Header. If MN has no security association, The MN sends the Mapping Refresh message to the CN to inform CN of the event that MN has moved. In this case, CN shall re-query the MA to obtain the new network prefix of MN. MN also sends Mapping Update Request message to the Mapping Agent (MA) to inform the current network prefix.

In order to track the current location of mobile node, LIN6 employs MA to maintain the mapping of the LIN6 ID and the network prefix, and makes use of the DNS to locate the MAs of MN. Each MA shall be assigned a predefined 64-bit value called MA IFID as the interface identifier. When MN is powered on and attaches to a network for the first time, it registers its current location with its MAs. When CN want to communicate with MN for the first time, the CN sends a query to the DNS sever and obtains the AAAA record, which consists of the network prefix of MA and the LIN6 ID of MN. CN generates the IPv6 address of MA by concatenating the upper 64 bits of the AAAA record and the MA IFID is used as the lower 64 bits of IP address, then it query the MN's MA the network prefix of MN. When MN moves to a new network, it registers the new network prefix with MA and CN by sending Mapping Update messages (MUs) with the Authentication Header or Mapping Refresh message.

Figure 6 shows the LIN6 network architecture and its operations.



1. Register MN Network Prefix->LIN6 ID at Bootstrap
2. Query MN FQDN
3. Response MA Network Prefix+LIN6 ID
4. Query MN LIN6 ID
5. Response MN Network Prefix
6. Establish Connection and Data Transfer
7. Move
8. Update the Mapping of LIN6 ID->MN Network Prefix
9. Go on Communication

Figure 6: LIN6 Mobility and its Operations

3.1.3. Analysis of Network Layer Mobility. Although micro-mobility protocols can improve MIP performance, MIP has still the weakness of inefficiency and complexity. MIPv6 has the advantages of inherent mobility and security support in IPv4 and routing optimization compared to MIPv4. However, Like MIPv4, MIPv6 has the same problem of third device, which increases failure probability of communication. And it has additional header overhead. The enhancements of HMIPv6, FMIPv6, and their combination improve the performance by eliminating signaling overhead, packet loss and handover latency etc, but their scalability and complexity is a concern.

In comparison with MIP, LIN6 is more fault tolerant because the HA in MIP can not be replicated to the subnet other than the home link while the MA introduced in LIN6 can be replicated anywhere on the Internet. And LIN6 has less overhead due to avoid the extension header and tunneling. That is, LIN6 does not use any packet interceptor or forwarder such as HA, so its routing is same as traditional IP-based routing. Conceptually, LIN6 adds a transient “presence” service to DNS lookup for dynamic locator mapping (from this sense, LIN6 can be also considered as introduction of a new layer), but it is only limited to IPv6.

Recently the MOBIKE WG has been working on adding features to IKEv2 to support mobility and multihoming called MOBIKE. MOBIKE handles IP address changes initiated by one of the endpoints of the security associations, and can keep the existing IKE and IPsec SAs in place without full re-keying. Note this is still under discussion and there is no concrete solution yet [14].

3.2. Mobility Support in Transport Layer

As the transport layer is subject to impact of mobility. A lot of work for performance improvement and mobility enhancement of TCP has been developed over past years [27-30]. Recently, the mobility support for the new transport layer protocols of SCTP and DCCP has been proposed. The basic idea of enabling transport layer mobility is to remove network layer dependences by using indirection, migration, tunneling or multi-homing etc. techniques.

3.2.1. TCP. There is much study to focus on the TCP as it is the most widely used transport layer. We classify the different proposals into two categories: Improving TCP performance in mobile Internet and Mobility support extension to TCP.

Improving TCP Performance in Mobile Internet

As traditional TCP can not adapt well to Internet mobility, a number of researchers aim to improve TCP performance in mobile Internet.

Indirect TCP (I-TCP) [27] and MTCP [28] focus on the BER problem of wireless link. In I-TCP and

MTCP, a TCP connection between MN and FN is split in two at a device called Mobile Support Station (MSS) and the connection between the MSS and MH is optimized for the wireless link. Both I-TCP and MTCP achieve better throughputs than standard TCP. Caceres and Iftode used the fast retransmission mechanism [29] to address the problem of short disconnections during handover. Haas developed an asymmetric transport layer protocol called Mobile-TCP [30] to minimize communication overhead on the MN. In Mobile-TCP, functions through algorithms and procedures are implemented with the lower complexity on MN than FN without sacrificing the performance and features.

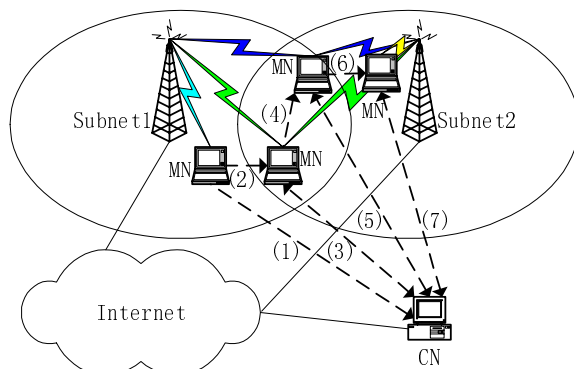
Mobility Extension to TCP

Other researchers consider the issue that how to maintain the ongoing connection when disconnection occurs due to the change of IP address.

Funato [31] developed a simple and secure redirection mechanism called TCP Redirection (TCP-R) to keep connections actively. The concept of TCP-R is to revise the pair of addresses in the ongoing TCP connections when the IP address associated to TCP connection is changed by TCP redirection options. In TCP-R, when MN initiates a new connection, it shall ascertain if CN is TCP-R aware or not, then may perform redirection operation. When MN moves and is assigned a new IP address, it sends a redirect message with RD_REQ option to CN. Upon CN receiving the message, it shall validate the connection authenticator with AT_REQ and AT_REP. If correct, it revises the pair of addresses of the ongoing TCP connection with the new MN’s IP address. Simultaneously, the MN also revises its own pair of IP addresses. Finally, they resume to communicating with the revised TCP connection. Snoeren and Balakrishnan [32] propose an end-to-end approach to support TCP mobility by migrating technique. TCP Migrate has the similar idea of TCP-R. It differentiates from TCP-R through different implementation by specifying different TCP migrate options. MSOCKS [33] presents an alternative TCP mobility support by split-proxy mechanism and extension to SOCKS [34]. In MSOCKS, when MN changes the IP address that a TCP connection uses to communication with the MSOCKS proxy, it shall open a new connection to the proxy and sends a RECONNECT messages with the connection identifier of the existing connection. Upon receiving a RECONNECT message, the proxy separates the old connection between MN and Proxy (MN-Proxy) from the connection between Proxy and CN (Proxy-CN), and concatenates in the new MN-Proxy connection. The proxy then concatenates the new connection to the Proxy-CN connection in place of the old MN-Proxy connection and closes the old connection. Once the concatenation is setup, the proxy sends an ok message to MN.

3.2.2. MSCTP. SCTP is a new IETF transport protocol. The feature of multi-homing provides an excellent platform for mobility support. In addition, ADDIP [35] extension enables SCTP mobility support called MSCTP [36].

In MSCTP, the MN initiates an SCTP association with the CN by negotiating a list of IP addresses. Among these addresses, one is chosen as the primary path for normal transmission, the other addresses are specified as active IP addresses. When MN reaches a new subnet and obtains a new IP address, MN sends Address Configuration Change (ASCONF) Chunk with Add IP Address parameter to inform the CN of the new IP address. On receiving the ASCONF, CN shall add the new IP address to the list of association addresses and reply the ASCONF-ACK Chunk to MN. While MN moving, MN may change the primary path to the new IP address by path management function. The SCTP association, therefore, can continue data transmission while moving to new network. MN also informs CN to delete the IP address of previous network from the address list by sending ASCONF Chunk with Delete IP Address parameter when MN confirms that the link of previous network has failed permanent. Figure 7 illustrates the operations of MSCTP.



1. Initiate an Association through Subnet1
2. Move
3. Add IP into the Association
4. Move
5. Change the Primary Path
6. Move
7. Delete IP from the Association

Figure 7: MSCTP Operations

3.2.3. DCCP. DCCP provides integrated mobility and multihoming support by defining DCCP-Move packet type, two new DCCP features of Mobility Capable feature and Mobility ID feature. DCCP specifies the mobility support optional and defaults to be off, so DCCP nodes must enable mobility support by Mobility Capable feature. Firstly, MN send a Change L option of Mobility Capable feature to inform CN that it would like to enable to changes its address during connection.

Then CN sends a Change R option to confirm MN. After that, MN sends a value of Mobility ID feature that will be used to identify connection. The value of Mobility ID feature is selected randomly for security reasons, also new value should be chosen after each move of MN. CN confirms value of Mobility ID feature by sending Conform L option. When MN reaches a new network and obtains the new IP address, it shall send a DCCP-Move packet containing Mobility ID value that was chosen for connection identification. Upon receiving DCCP-Move packet, CN shall send DCCP-Sync message to MN, and change its connection state and start using new address of MN.

3.2.4. Analysis of Transport Layer Mobility. The extensions to TCP proposed for improving transport performance in mobile Internet can not deal with mobility well on their own. Their main purpose is merely to minimize degradation of transport performance. The mobility enhancements of TCP-R, TCP Migrate and MSOCKS to TCP can handle mobility and keep all features of the standard TCP. Their operations are done in a secure way.

MSCTP provides an alternative solution at the transport layer. It can support seamless handover and improve transport performance. Unlike TCP, SCTP uses four-step negotiation process to initiate an association, which can prevent the denial of service attacks such as SYN attacks, and IPsec is used to secure the SCTP communication. During the Addition/Deletion of IP addresses of mobility, MSCTP suggests using the IP AH to protect the existing association from being hijacked or attacked. However, the current MSCTP proposal only illustrates the basic requirements for Internet mobility. Some essential issues, such as when and by which criteria the primary path to be changed or the addition and deletion of the IP addresses mapped to the SCTP association should occur during handover, are yet left for further study. Moreover, MSCTP by itself does not handle support location management, thus a proposal on reusing MIP for location management in MSCTP is proposed in [37].

Similarly, the current specification of DCCP is at its primitive stage. There are many problems unsolved. For example, DCCP has no support for simultaneous movements of both communicating endpoints, i.e. DCCP supports mobility of only one endpoint, and the other one should remain stationary.

3.3. Providing Mobility Support in a New Layer

Traditional TCP/IP protocols are already rather heavily loaded with functionalities added over the years, the optimization and adding new functionalities to support mobility is very difficult. Therefore, another idea for Internet mobility is to introduce a new layer such as

HIP, MASK and MOBIKE where the Internet mobility is deployed.

3.3.1. HIP. HIP [38] is designed to establish secure communication and to provide continuity of communication. Similar to LIN6, HIP is based on the idea of separating location from identity by an interposed host identity protocol layer that operates between network and transport layers (see Fig. 8). HIP introduces a new Host Identity namespace called Host Identifier (HI). The transport layer connection (Socket) is bound to HI instead of IP address, and IP address becomes pure routing messages. The HI is dynamic mapped to one or more IP addresses at the HIP layer.

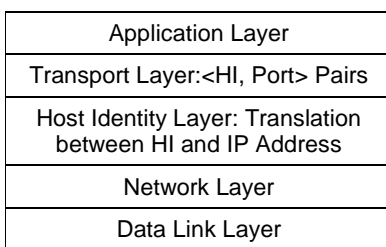


Figure 8: New Proposed TCP/IP Structure

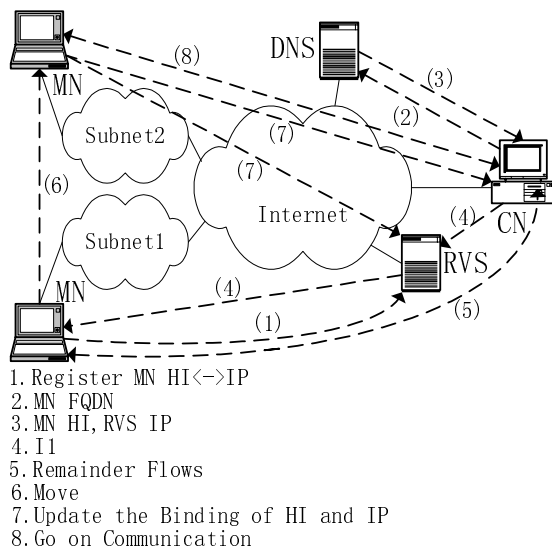


Figure 9: HIP Mobility and its Operations

In HIP, the dynamic binding between HI and IP address is achieved by using the Update packet with HIP Readdress Packets (REA) parameter. Besides, HIP employs the Rendezvous Sever (RVS) to provide location management. On HIP initiation, the initiator retrieves the RVS IP address by looking up the domain name of the peer from DNS with the HIPRVS RR, and sends I1 with destination HIT packet to the RVS. The RVS shall forward the initial HIP packet to the peer at its current location. Afterward, after receiving I1, the peer will complete the HIP initiation directly without the help of RVS. When MN moves while

communication is ongoing and gets a new IP address, MN shall send a HIP Update packet with REA to inform CN of the new IP address, and CN shall respond with a HIP Update with ACK. For security concern, CN may verify that the MN is reachable through the new IP address. Once CN verifies successfully, it makes the new IP address as active and removes the old address, then CN can communicate with the new IP address. Figure 9 illustrate the operations of HIP.

3.3.2. MAST. MAST is proposed by Crocker [5] for Internet mobility and multi-homing. Like HIP, MAST defines a layer between network and transport layers but avoiding creating new namespace by using the existing IP addresses. It only maps different IP addresses to a single IP address that may be initially IP address assigned for the association.

MAST defines a mechanism that supports association of multiple IP addresses with any transport association. The MAST association is manipulated with request/response messages, which are used to initially establish the MAST association, to update the set of valid IP addresses, to query association status, to convey error information and to terminate the association etc.

With MAST, when the MN moves across the Internet, the IP addresses of MN locators may be added and removed, while the initially IP address of MN identifier continue to be bound to transport layer and other addresses of MN locators are mapped to that IP address of identifier by MAST control exchange. Over the life of a mobile transport association, different addresses of MN locators might be active at different times.

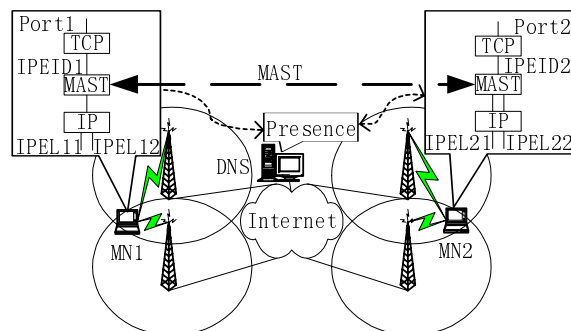


Figure 10: MAST Architecture

To find mobile targets, MAST uses DNS to provide the information of dynamic presence service relating to MN. The DNS SRV record is defined to reference a dynamic presence service through which an endpoint can register its current set of IP addresses. Besides, MAST also specifies that the MN registers its current address with dynamic the presence service available through the Extensible Messaging and Presence

Protocol (XMPP) [42]. Figure 10 illustrate the architecture of MAST.

Security considerations: To resist the attacks of hijacking an association, MAST uses association-specific weak authentication [43], which ensures that later packets come from the same source as initial packet. Besides, IPsec or TLS is also suggested for other security issues like spoofing and redirection etc. In essence, MAST may be no worse than the current generally used level of security, but it does not increase the security either.

3.3.3. Analysis of New Layer Mobility. HIP supports multihoming by dynamic mapping from one Host Identifiers to multiple IP addresses. And it resolves the problem of simultaneous movement of endpoints by resending the HIP readdress message to the RVS if no reply is received. However, RVS changes the basic property by replacing the IP addresses of their client nodes in the DNS with their own. The IP addresses in DNS entry hence no longer directly designate interfaces of an endpoint. And it suffers from failures because the II packet during initializing a connection must be relayed by RVS. Besides, the applications that have followed the structure of old layers have to be modified to it.

MAST suggests the same core MAST does not define any new namespace or addressing structure and requires no change to the Internet infrastructure with supporting both existing IPv4 and IPv6, no change to IP modules or transport modules in the end-systems, and no new administrative effort. And it has no additional packet header overhead and minimal additional packet-processing overhead. Hence MAST has a low barrier to adoption and use, while permitting more advanced functions with more extensive adoption and modification. However as primitive protocol, there are many open issues to be resolved. For example, the optimal locator selection is largely an issue.

3.4. Mobility Support in Application Layer

In the context of application layer, there are also attempts to support Internet mobility. This subsection presents Internet mobility support for SIP and DDNS in application layer.

3.4.1. SIP. SIP was initially developed by IETF as an application layer multimedia signaling protocol, but it has the potential capabilities for Internet mobility. In SIP, When the MN initiates a session with the CN, it sends an INVITE message and the normal SIP signaling procedure is performed to establish the session. When the MN accesses to a new network and obtains a new IP address while the session is ongoing, MN shall send RE-INVITE message maintaining the same Call-ID of the existing session but replacing the Contact and c fields in the SIP and SDP headers with

the new IP address to CN. And the UA sends a REGISTER message to the home SIP server to update the location information stored there.

3.4.2. DDNS. As mentioned in section 2, traditional DNS is restricted in mobile Internet. To resolve the problem, P. Vixie et al. [11] propose a method for dynamic updating RRs or RRsets from a specified zone by specifying the UPDATE messages. Because most applications ubiquitously resolve node name to an IP address at the beginning of communication, DDNS can be considered for the location management in the mobile scenario where the MN acts as a server and other nodes actively originate communication with the MN.

To locate the MN as it moves to new network, the MN dynamically registers and updates its name-to-address or FQDN-to-IP entry (A record) with new IP address to DNS servers by sending DNS UPDATE messages. Then whenever a CN wants to communicate with the MN, it will query DNS sever with FQDN of the MN, and DNS sever shall response with the current IP address of MN. Finally, the CN can initiate and establish communication with the MN directly. Figure 11 illustrate the location management of DDNS in mobile Internet.

Security considerations: The dynamic UPDATE messages are based on authenticated requests [44] and transactions is used to provide authorization by TSIG [45] or SIG (0) [46, 47]. Only authorized sources are allowed to make changes to a zone's contents.

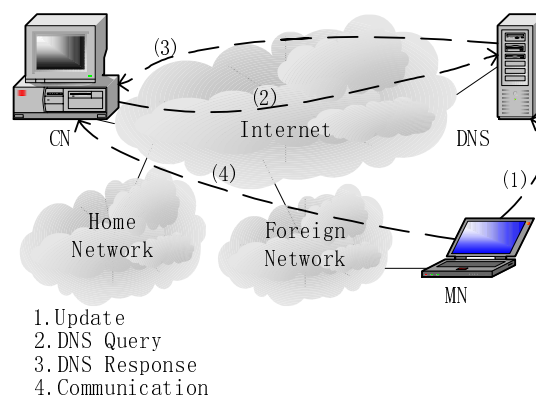


Figure 11: DDNS Location Management

3.4.2. Analysis of Application Layer Mobility. SIP provides Internet mobility support without any modifications of lower layer protocols, but the handover of SIP incurs considerable handover latency that is not suitable to real-time communications [39-40]. To improving SIP mobility, Dutta [38] optimizes SIP mobility management by using the intra-domain solution, which limits the movement indication to within the domain, to reduce handover latency and minimize packet loss. Kim et al. [41] proposes a

mechanism of Predictive Address Reservation with SIP, which reduces handover latency by proactively processing the address allocation and session update using link layer information of wireless networks.

DDNS utilize existed DNS for location management, would not require special servers as MIP. However, The DNS registration delay needs to optimize. Besides, as DDNS can not maintain ongoing communication in the mobile Internet, it is used as a candidate approaches for location management along with other solutions.

4. Comparison of Different Paradigms for Internet Mobility Support

In this section we will qualitatively evaluate the mobility solutions on the layer category level summarized above from two aspects of functional requirements and performance metrics. We emphasize the comparison is not complete for solutions in question but the main factors are found with the sort of comparing approach.

Table1. Functions of Proposed Paradigms: A Comparison

Paradigm	Ext. netw. layer			Ext. transport layer			Intro. a new layer		Ext. appl. layer	
	MIP	LIN6	MOBIKE	TCP	SCTP	DCCP	HIP	MAST	SIP	DDNS
Handover	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Location	✓	✓	✓				✓	✓	✓	✓
Multi-homing			✓		✓	✓	✓	✓	✓	✓
Applications	✓	✓		✓	✓	✓				✓
Security		✓	✓		✓	✓	✓	✓	✓	✓

Table2. Required Changes to Existing Systems: A Comparison

Paradigm	Ext. netw. layer			Ext. transport layer			Intro. a new layer		Ext. appl. layer	
	MIP	LIN6	MOBIKE	TCP	SCTP	DCCP	HIP	MAST	SIP	DDNS
Host	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Router	✓	✓							✓	
Third Device	✓	✓	✓				✓	✓	✓	
TCP/IP Layering							✓	✓		

4.2. Performance Comparison

Following analysis compares the performance of different solutions based on metrics of handover latency, packet loss, signaling overhead and throughput. The handover mechanism of network layer suffers from large handover latency and considerable packet loss caused proxy and no support of multihoming, although many techniques such as anticipate etc. have been developed to address the problems. Advantages to transport layer mobility include inherent route optimization, no dependence on the third device make it the possibility of seamless handover with multihoming support, and minimization of packet loss with the ability to pause transmissions in expectation of a mobility-induced temporary disconnection.

Mobility solution at the network layer also involves signaling overhead problem caused by tunneling and extension headers etc. Transport layer solution seems to alleviate the problem because they

4.1. Functional Requirement Comparison

Firstly, we summarize and compare the mobility support solutions based on requirements for handover management, location management, multi-homing, applications and security. Table 1 summarizes the how the requirements presented in Section 2 are supported by the solutions presented above. From the table, we can conclude that none of these solutions fulfill all requirements. Network layer does not yet support multi-homing. New layer solution of HIP must define new API for the HI, which requires modification of current applications. Transport layer by itself can not track node, so it is short of location management function. They depend on other layers for location management. For example, if dynamic DNS is employed, it may take quite some time to globally converge to a host's current address, by which time it may be ready to move again. Application layer solutions are only appropriate for specifically applications such as real-time multimedia.

manage mobility by negotiating and switching connections directly between endpoints.

Besides, network layer by itself can not guarantee that the efficiency of transport connections is maintained and can not handle the degradation of throughput caused by congestion control. Transport layer mobility improves degradation of throughput effectively by implementing policies that reset congestion control after reattachment.

4.3. Comparison of Required Changes to the TCP/IP Stack

In order to maintain backward compatibility, the network and protocol infrastructure concern is another important factor in deployment including required changes to endpoint and intermediate router or addition of third entity such as proxy, agent etc. for network infrastructure, as well as change to protocol infrastructure. Table 2 illustrates the required changes comparison for different solutions. Network layer

solutions are based on routing mechanism, so they require changes to endpoint and router for addressing binding. In additions, they need third device of agents for packet forwarding and location management. Because transport layer solutions are based on end to end model, they require no change to intermediate routers. Besides, they are absent from location management by themselves, there is not deployment of third device. Therefore, the transport layer solutions require very little infrastructure change. New layer solutions need modification of endpoint. And it employs RVS for location management, so it also needs addition of third device. Besides, the introduction of new protocol layer also destroys the traditional TCP/IP infrastructure. Similarly, the application solution of SIP employ proxy server to relay flows and register server to locate MN, it need add third device and change of endpoint.

5. Conclusion

In this paper, we analyzed the problems of traditional TCP/IP stack caused by the mobility of nodes and their wireless links, illustrated the Internet mobility issue may be affected many layers, and illustrated many layers of TCP/IP stack have the negative effect on the Internet mobility issue. Many solutions have been proposed solve the operability problem of integrating mobile networks into the Internet. Now, we need to have complete solution, which should be designed with the problems of mobility and wireless connections in mind.

We presented a survey of different mobility support paradigms for the Internet. From our comparisons and the discussion of the advantages and disadvantages of each paradigm, we concluded that current mobility solutions do not solve all general problems related to Internet mobility and it is hard to dictate which one is most suitable: Individual layers contribute to Internet mobility, while the technology is important, the market will decide. Link layer mobility support is foundational in mobile Internet, but it constrains within a limited domain and can not preserve higher layer connections. Although the network layer solutions can handle most of requirements, it has slowly deployment in practice for ineffective and complexity. Transport layer solutions can fulfill handover management efficiently, but it is short of the ability of location management by itself. And application layer approaches are restricted in specific applications.

To provide an effective solution with the issues of both basic functional and performance requirements for Internet mobility support in mind, we conclude with the features that need to be satisfied in the mobile Internet:

1. Can efficiently deal with handover. For example, using anticipating technique of radio trigger etc. to

detect handover and perform routing/path update and location registration process in advance.

2. Can handle various mobile scenarios of the endpoints, including client-server scenario where the MN only originates the sessions and the point-to-point scenario where the sessions may be originated at either one endpoint of communicating peers, by enhance location management such as DDNS etc.

3. Provide end-to-end mobility and avoid third party entities or tunneling mechanism which improve the complexity and reduce the performance of mobility.

4. Take advantage of the multihoming, which can make for seamless handover and improving performance of mobility with its redundancy and load share etc. features simultaneously.

5. Avoid erroneously triggering congestion control mechanisms, arising from the wireless link characteristics of lossy and bursty high BER etc, at the transport layer by enhancing signaling mechanism between link and transport layers.

6. Preferably provide compatibility. That is, do not require change or impact in applications, network architecture, TCP/IP structure, or entities.

7. Take into consideration of the security in mobile Internet.

The efficient Internet mobility management is a more challenging issue. In order to satisfy these features recommended above, it needs all the layers' participation in a highly cooperative way. Therefore, we anticipate a multi-layer architecture for advanced mobility support and we suggest the transport layer as the main candidate assisted with other layers together for Internet mobility support.

References

- [1] T.R. Henderson, "Host mobility for IP networks: a comparison", IEEE Network, Nov. 2003, pp. 18-26.
- [2] W.M. Eddy, "At what layer does mobility belong?", IEEE Communications Magazine, Oct. 2004, pp. 155-159..
- [3] K. Kuladinithi, A. Kongseng, and S. Aust et al., "Mobility management for an integrated network platform, Mobile and Wireless Communications Network", 4th International Workshop, Sept. 2002, pp. 621-625.
- [4] Hala Elaarag, "Improving TCP performance over mobile networks", ACM Computing Surveys, Sep. 2002, pp. 357-374.
- [5] C. Perkins. "IP Mobility Support for IPv4", RFC 3344, Aug. 2002.
- [6] D. Johnson, C. Perkins, and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.
- [7] F. Teraoka, M. Ishiyama, and M. Kunishi, "LIN6: A Solution to Multihoming and Mobility in IPv6", draft-teraoka-multi6-lin6-00.txt, December 2003.
- [8] R. Stewart, Q. Xie, and K. Morneault, "Stream Control Transmission Protocol", RFC 2960, Oct. 2000.
- [9] Eddie, and Kohler, "Datagram Congestion Control Protocol Mobility and Multihoming", draft-kohler-dccp-mobility-00.txt, July 2004.

- [10] F. Vakil, A. Dutta, and J-C. Chen, "Supporting Mobility for Multimedia with SIP", draft-itsumo-sipping-mobility-multimedia-01.txt, July 2001.
- [11] P. Vixie, S. Thomson, Y. Rekhter, et al., "Dynamic Updates in the Domain Name System (DNS UPDATE)", RFC 2136, April 1997.
- [12] P. Nikander, J. Arkko, and T. Henderson, "End-Host Mobility and Multi-Homing with Host Identity Protocol", draft-nikander-hip-mm-02.txt, July, 2004.
- [13] D. Crocker, "Multiple Address Service for Transport (MAST): an Extended Proposal", draft-crocker-mast-proposal-01.txt, Sep. 2003.
- [14] T. Kivinen and H. Tschofenig, "Design of the MOBIKE Protocol", draft-ietf-mobike-design-01.txt, December 2004.
- [15] C. Perkins, "IP Encapsulation within IP", RFC 2003, Oct. 1996.
- [16] B. Aboba, "IAB Considerations for the Split of Identifiers and Locators", draft-iab-id-locsplit-00.txt, Mar. 2004.
- [17] R. Droms, "Dynamic Host Configuration Protocol", RFC2131, March 1997.
- [18] C. Perkins, "Mobile IP Regional Registration", draft-ietf-mobileip-reg-tunnel-09.txt, 2004.
- [19] K. El Malki, H. Soliman, "Fast Handoffs in Mobile IPv4", draft-elmalki-mobileip-fast-handoffs-03.txt, 2000.
- [20] R. Ramjee, T. La Porta, S. Thuel, et al. "HAWAII: A Domain-based Approach for Supporting Mobility in Wide-area Wireless Networks", Proc. IEEE International Conference on Network Protocols, 1999.
- [21] A. Valko, "Cellular IP: A New Approach to Internet Host Mobility", ACM SIGCOMM Computer Communication Review, Jan. 1999, pp. 50-65.
- [22] R. Droms, J. Bound, B. Volz, et al., "IPv6 Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.
- [23] S. Thomson and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, Dec. 1998.
- [24] H. Soliman, C. Castelluccia, K. El-Malki et al., "Hierarchical MIPv6 mobility management (HMIPv6)", draft-ietf-mobileip-hmipv6-08.txt, June 2003.
- [25] G. Tsirtsis, A. Yegin, C. Perkins et al., "Fast Handovers for Mobile IPv6", draft-ietf-mobileip-fast-mipv6-08.txt, Oct. 2003.
- [26] H.Y. Jung, S.J. Koh, H. Soliman et al, "Fast Handover for Hierarchical MIPv6 (F-HMIPv6)", draft-jung-mobileip-fastho-hmipv6-04.txt, June 2004.
- [27] A. Bakre and B. R. Badrinath, "I-TCP: Indirect TCP for Mobile Hosts", Proceedings of the 15th International Conference on Distributed Computing Systems, Vancouver, Canada, June 1995, pp. 136-143.
- [28] R. Yavatkar and N. Bhagawat, "Improving End-to-End Performance of TCP over Mobile Internetworks", IEEE Workshop on Mobile Computing Systems and Applications, Santa Cruz, CA, 1994.
- [29] R. Caceres., L. Iftode, "Improving the performance of reliable transport protocols in mobile computing environments", IEEE Journal on Selected Areas in Communications, 1995, pp. 850-857.
- [30] Z.J. Haas, "Mobile-TCP: An Asymmetric Transport Protocol Design for Mobile Systems", IEEE International Conference on Communications (ICC'97), Montreal, Canada, 1997.
- [31] D. Funato, K. Yasuda, and H. Tokuda. "TCP-R: TCP mobility support for continuous operation", Proc. IEEE ICNP'97, 1997, pp. 229-236.
- [32] A. C. Snoeren and H. Balakrishnan, "An End-to-End Approach to Host Mobility", Proc. of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking, Aug. 2000.
- [33] David A. Maltz and Pravin Bhagwat, "MSOCKS: An Architecture for Transport Layer Mobility", Proceedings of IEEE Infocom, 1998
- [34] M. Leech, M. Ganis, Y. Lee, et al., "SOCKS protocol version 5", RFC1928, April 1996.
- [35] R. Stewart, "Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration", draft-ietf-tsvwg-addip-sctp-09, June 2004.
- [36] M. Riegel, and M. Tuexen, "Mobile SCTP", draft-riegel-tuexen-mobile-sctp-03, Oct. 2004.
- [37] S.J. Koh and Qiaobing Xie, "Mobile SCTP with Mobile IP for Transport Layer Mobility", draft-sjkoh-mobile-sctp-mobileip-04.txt, June 2004.
- [38] P. Nikander, J. Ylitalo, and J. Wall, "Integrating Security, Mobility, and Multi-Homing in a HIP Way", Proceedings of Network and Distributed Systems Security Symposium (NDSS'03), San Diego, CA, Feb. 2003, pp. 87-99.
- [39] E. Wedlund, H. Schulzrinne, "Mobility Support using SIP", IEEE/ACM Multimedia conference WOWMOM, Aug. 1999, pp.76-82.
- [40] A. Dutta, et al., "Implementing a Testbed for Mobile Multimedia", GLOBECOM '01, IEEE, Nov. 2001, pp.25-29.
- [41] W. Kim, M. Kim, K. Lee, C. Yu and B. L. Link. "Layer Assisted Mobility Support Using SIP for Real-time Multimedia Communications", ACM MobiWac, 2004.
- [42] P. Saint-Andre, J. Miller, "Extensible Messaging and Presence Protocol (XMPP): Core", draft-ietf-xmpp-core-24, May 6, 2004.
- [43] J. Arkko and P. Nikander, "Weak Authentication: How to Authenticate Unknown Principals without Trusted Parties", Proceedings of Security Protocols Workshop 2002, Cambridge, UK, April 16-19, 2002, pp. 5-19.
- [44] B. Wellington, "Secure Domain Name System (DNS) Dynamic", RFC 3007, Nov. 2000.
- [45] P. Vixie, O. Gudmundsson, and D. Eastlake et al., "Secret Key Transaction Authentication for DNS (TSIG)", RFC 2845, May 2000.
- [46] D. Eastlake, "DNS Request and Transaction Signatures (SIG (0) s)", RFC 2931, September 2000.
- [47] D. Eastlake, "Domain Name System Security Extensions", RFC 2535, March 1999.