Wireless Personal Communications (2007) 43:589–603
DOI 10.1007/s11277-007-9251-x

An Empirical Analysis of Handoff Performance for SIP, Mobile IP, and SCTP Protocols

SHERALI ZEADALLY[1] and FARHAN SIDDIQUI[2]

[1]*Network Systems Laboratory, Department of Computer Science and Information Technology, University of the District of Columbia, Washington DC 20008, USA*
E-mail: Zeadally@cs.wayne.edu
[2]*Department of Computer Science, Wayne State University, Detroit, MI 48202, USA*

Abstract. Over the last decade, we have witnessed a growing interest in the design and deployment of various network architectures and protocols aimed at supporting mobile users as they move across different types of networks. One of the goals of these emerging network solutions is to provide uninterrupted, seamless connectivity to mobile users giving them the ability to access information anywhere, anytime. Handoff management, an important component of mobility management, is crucial in enabling such seamless mobility across heterogeneous network infrastructures. In this work, we investigate the handoff performance of three of the most widely used mobility protocols namely, Mobile IP, Session Initiation Protocol (SIP), and Stream Control Transmission Protocol (SCTP). Our empirical handoff tests were executed on an actual heterogeneous network testbed consisting of wired, wireless local area, and cellular networks using performance metrics such as handoff delay and handoff signaling time. Our empirical results reveal that Mobile IP yields the highest handoff delay among the three mobility protocols. In addition, we also found that SIP and SCTP yield 33 and 55% lower handoff delays respectively compared to Mobile IP.

Keywords: handoff, protocol, performance, network,wireless.

1. Introduction

The demand for ubiquitous information access has led to the convergence of several types of networks including Ethernet Local Area Network (LAN), General Packet Radio Service (GPRS), Global System for Mobile Communication (GSM), Wireless Local Area Network (WLAN), Bluetooth, etc. In such heterogeneous environments mobility management is the basis for providing continuous network connectivity to mobile users roaming between these access networks. There are two major components of mobility management: Location management and Handoff management. Location management enables the network to discover the current attachment point of the mobile user. Handoff management enables the Mobile Node (MN) to maintain the network connection as it continues to move and change its access points or base stations to the network.

Several protocols have been proposed [4, 5, 8] to address the issue of mobility management in heterogeneous networks. These approaches operate at different levels of the network protocol stack.

- ***Network Layer*** **:** Mobile IP [21] was proposed by the Internet Engineering Task Force (IETF) to handle mobility management at the network layer. It handles mobility by redirecting packets from a MN's home network to the MN's current location. Deployment of

Mobile IP requires network servers including a home agent and a foreign agent that are used to bind the home address of a MN to the care-of address at the visited network and provide packet forwarding when the MN is moving between IP subnets. It is worthwhile noting that Mobile IP refers to Mobile IP version 4 throughout this paper.

- ***Application Layer*** **:** The Session Initiation Protocol (SIP) [23] is an application layer protocol that keeps mobility support independent of the underlying access technologies. With the SIP approach, when an MN moves during an active session into a different network, it first receives a new network address, and then sends a new session invitation to the Correspondent Node (CN) Subsequent data packets from the CN are forwarded to the MN using the new address. The MN also needs to register its new IP address with a SIP server called a Registrar to enable other nodes on the network to reach it by querying the Registrar server.
- ***Transport Layer*** **:** A third approach for mobility management has been proposed that operates at the transport layer. The Stream Control Transmission Protocol (SCTP) [29] uses this approach. The SCTP-based approach uses multihoming to implement mobility management. The multihoming feature allows SCTP to maintain multiple IP addresses. Among those addresses, one address is used as the primary address for the current transmission and reception. Other addresses (secondary) can be used for retransmissions. The multihoming feature of SCTP provides the basis for mobility support since it allows a MN to add a new IP address, while holding an old IP address already assigned to it.

In this paper we present an empirical analysis of handoff performance for SIP, Mobile IP and SCTP protocols running over different types of wired/wireless access networks. The rest of this paper is organized as follows. We present related works and contributions of this paper in Section 2. In Section 3 we give an overview of the three mobility management protocols: SIP, Mobile IP, and SCTP. Section 4 presents the experimental procedures and testbed setup used to conduct our performance evaluation tests. In Section 5 we present an analysis of handoff performance results. Finally, in Section 6 we make some concluding remarks and we discuss future work in Section 7.

2. Related Works and Contributions

2.1. Related Works

In the past few years, many researchers have investigated the performance of mobility protocols such as Mobile IP, SIP, and SCTP over wired and wireless networks.

In the area of Mobile IP performance, Hernandez and Helal [14] identified the limitations of Mobile-IP in terms of throughput, handoff and packet loss of a train moving at different velocities and the effect of different base station interleaving distances on throughput and packet loss. Saleh [24] investigated the performance of Mobile IP in the context of an interworking architecture between 802.11 WLAN and 2.5/3G CDMA cellular networks. Zhang et al. [34] proposed a mailbox-based approach that combines the benefits of approaches such as mobile IP route optimization and local registration to achieve adaptive location management.

The study of SIP performance has also received a lot of attention recently. Yeh et al. [33] discussed the implementation of SIP terminal mobility and presented a performance evaluation of SIP user agents developed with open-source libraries. They measured the delay involved using SIP mobility in both, IPv4 and IPv6 environments. They found that SIP mobility is suitable for supporting seamless handoffs for VoIP communications. Gokhale and Lu [13] demonstrated the feasibility of using SIP-based APIs in heterogeneous network infrastructures to measure the signaling performance of VoIP using SIP. Fathi et al. [9] presented an empirical evaluation of the session setup delay of SIP. They concluded that SIP-over-UDP makes the session setup delay 30% less than SIP-over-TCP. They also found that the SIP setup delay is much lower than the H.323 setup delay. Wu et al. [32] presented an analysis of the delay associated with vertical handoff using SIP in the WLAN-UMTS internetworking environments. Their analytical results showed that WLAN-to-UMTS handoff incurs unacceptable delay for supporting real-time multimedia services, and is mainly due to transmission of SIP signaling messages over erroneous and bandwidth- limited wireless links.

In the case of SCTP, several performance-related issues have been explored by various researchers. Shi et al. [26] presented a performance evaluation of SCTP in wireless environments and they found that SCTP-multihoming can provide better throughput performance and more robustness in wireless multi-access scenarios. Liu et al. [18] proposed a new approach to improve the performance of SCTP in wired-wireless environments by avoiding unnecessary congestion window decreases. Fracchia et al. [1] proposed a modification to the SCTP protocol to support the selection of the best available path based on available bandwidth and packet losses. Funasaka et al. [2] proposed a new path switching strategy for SCTP to improve the switching delay performance.

2.2. Contributions of this Work

Park and Dadej [20] have also investigated the performance of various IP mobility architectures. They provided recommendations for using mobile IP and related optimization mechanisms for selected wireless Internet applications. However, their performance results (similar to many other previous published results in this area) were based entirely on simulation tests using the OPNET modeling environment. In contrast, one of the goals of this work was to investigate handoff performance by exploiting *practical* performance measurement tests (under typical network conditions) over real heterogeneous wired/wireless networks. By performing our measurement tests over a real network testbed, we have a deeper insight into the actual performance delivered by actual systems and networks (compared to performance analyses done using only simulation which often makes it hard to capture practical constraints and limitations of real networks). Based on the early discussions presented on related works, we conclude that the performance of each of the mobility protocol (SIP, SCTP, Mobile IP) has indeed been *extensively* analyzed and studied by many other researchers. However, little work has been done (to the best of our knowledge) on side-by-side performance comparisons of SIP, SCTP, and Mobile IP. To address this issue, in this paper, we present a side-by-side empirical evaluation of handoff performance of all three mobility protocols when running over an actual heterogeneous network composed of wired and wireless networks. This is another major contribution of this work.

3. Mobility Protocols

3.1. Mobility With SIP

Session Initiation Protocol is an application-layer control protocol that can establish, modify, and terminate multimedia sessions [23]. SIP defines several logical entities, namely user agents, redirect servers, proxy servers, and registrars. SIP inherently supports personal mobility and can be extended to support service and terminal mobility [25]. Terminal mobility allows a device to move between IP sub-nets, while continuing to be reachable for incoming requests and maintaining sessions across subnet changes. Mobility of hosts in heterogeneous networks is managed by using the terminal mobility support of SIP.

Terminal mobility requires SIP to establish a connection either during the start of a new session, when the terminal or MN has already moved to a different location, or in the middle of a session. The former situation is referred to as *pre-call mobility*, the latter as *mid-call* or *in-session mobility*. For pre-call mobility, the MN re-registers its new IP address (acquired via DHCP in the case of WLAN and Ethernet networks and from the Gateway GPRA Support Node in the case of GPRS networks) with the Registrar server by sending a REGISTER message, while for mid-call mobility the terminal needs to notify the CN or the host communicating with the MN by sending a re-INVITE message about the terminal's new IP address and updated session parameters. The CN starts sending data to the new location as soon as it receives the re-INVITE message. The MN also needs to register with the redirect server in the home network for future calls. Figure 1 shows the messages exchanged for setting up a session between a MN and a CN and continuing it after changing the access network.

3.1.1. SIP Protocol Structure

Session Initiation Protocol [23] is structured as a layered protocol, which means that its behavior can be described in terms of a set of fairly independent processing stages with only a loose coupling between each stage. The lowest layer of SIP deals with the syntax and encoding of SIP information. The encoding is specified using an augmented Backus-Naur Form grammar (BNF) [3]. The second layer is the transport layer. It defines how a client sends requests and receives responses and how a server receives requests and sends responses over the network. The third layer is the transaction layer. Transactions constitute a fundamental component of SIP. A transaction is a request sent by a client (using the transport layer) to a server, along with all responses corresponding to that request sent from the server back to the client. The transaction layer handles application-layer retransmissions, matching of responses to requests, and application-layer timeouts. The transaction layer has a client component (referred to as a client transaction) and a server component (referred to as a server transaction). The layer above the transaction layer is called the Transaction User (TU). When a TU wishes to send a request, it creates a client transaction instance and passes the request along with the destination IP address, and port to which to send the request. A TU that creates a client transaction can also cancel it. When a client cancels a transaction, it requests the server to stop further processing, and to revert to the state that existed before the transaction was initiated, and generate a specific error response corresponding to that transaction. This is done with a CANCEL request.

3.2. Mobility with Mobile IP

Mobile IP is a mobility management protocol proposed to solve the problem of node mobility by redirecting packets to the MN's current location. The Mobile IP architecture is shown in Figure 2. Its main components include a Home Agent (HA) and a Foreign Agent (FA). HA is a router on a MN's home network, which encapsulates datagrams for delivery to the MN when it is away from home, and maintains current location information for the MN. FA is a router on a MN's visited network (foreign network) that provides routing services to the MN when

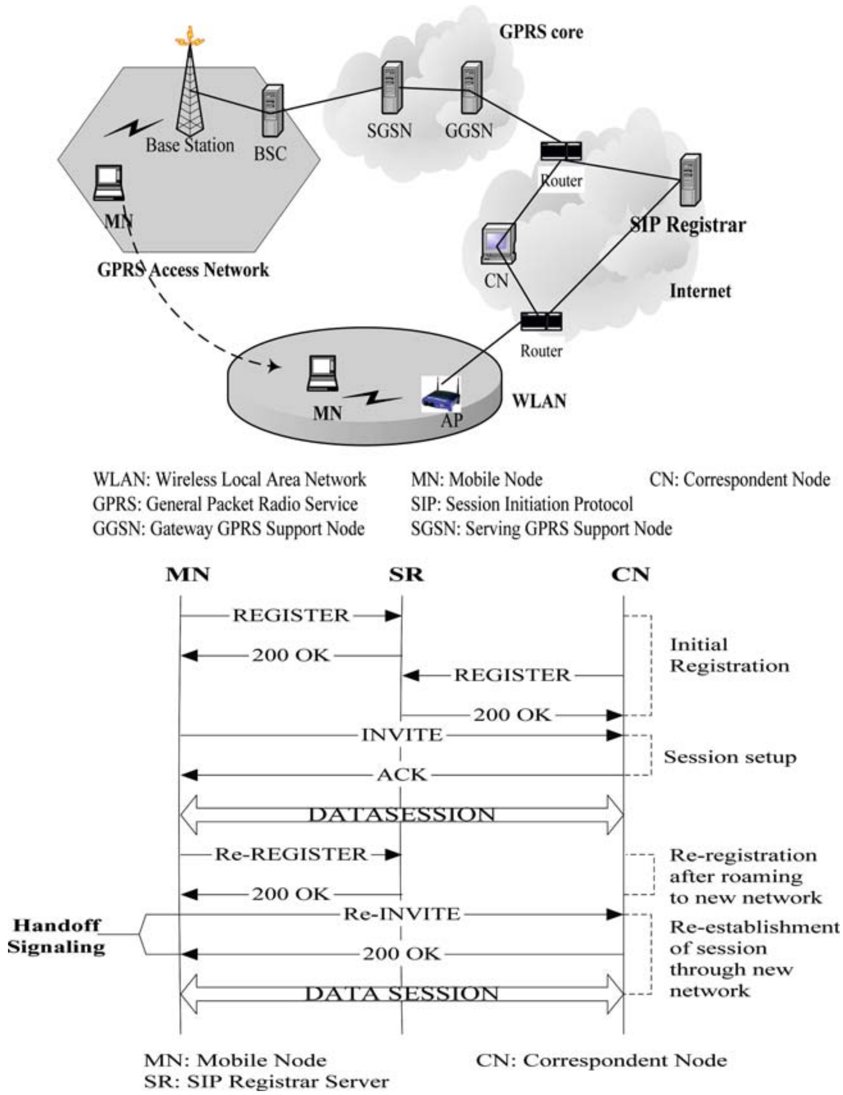

Figure 1. SIP-based mobility management.

registered. The FA decapsulates and delivers datagrams, tunneled by the MN's HA to the MN. When a MN moves out of its home network it must obtain another IP. So, in Mobile IP, a mobile host uses two IP addresses: a fixed home address (a permanent IP address assigned to the host's network) and a *care-of-address* – a temporary address from the new network (i.e., foreign network) that changes at each new point of attachment [5]. When the MN moves, it needs first discover its new care-of-address. The care-of-address can be obtained by periodic advertising from the FA through broadcasting. The MN then registers its care-of-address with its home agent by sending a Registration Request to its home agent via the foreign agent. The HA then sends a Registration Reply either granting or denying the request. If the registration process is successful, any packets destined for the MN are intercepted by the HA, which encapsulates the packets and tunnels them to the FA where decapsulation takes place and the packets are then forwarded to the appropriate MN.

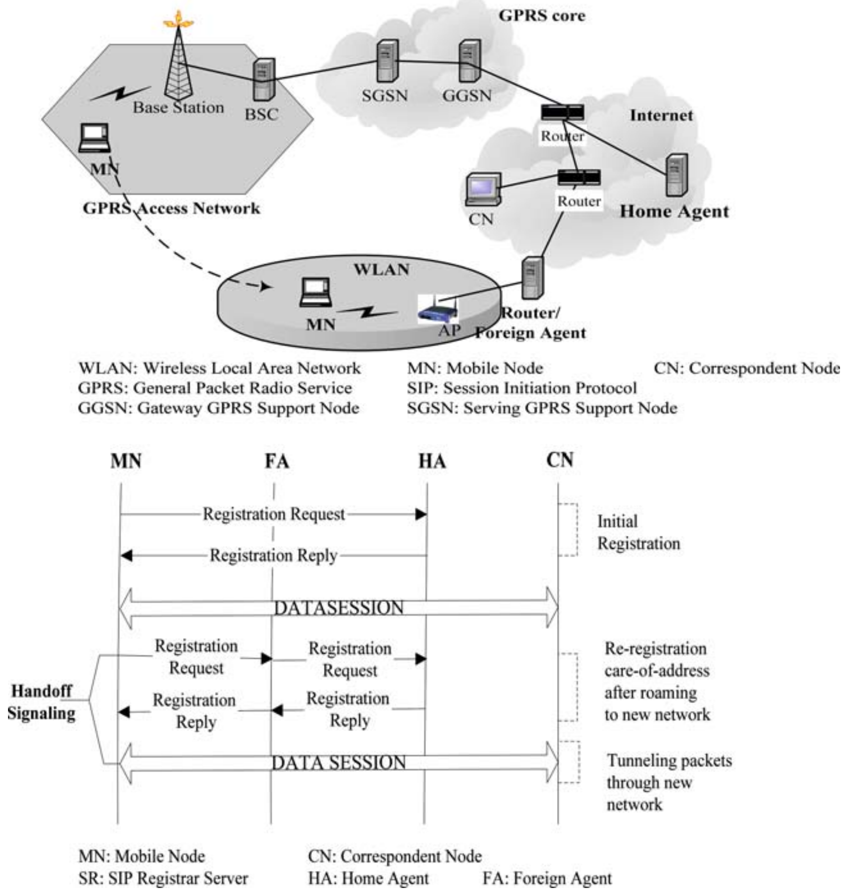

Figure 2. Mobile-IP-based mobility management.

3.3. Mobility with Stream Control Transmission Protocol (SCTP)

The SCTP [29] is a reliable, connection-oriented transport protocol that operates over a potentially unreliable connectionless packet service, such as IP. Before peer SCTP users can send data to each other, a connection must be established between two endpoints. This connection is called an association in SCTP context. A cookie mechanism is employed during the initialization of an association to provide protection against security attacks. Figure 3 shows a sample SCTP message flow. An essential property of SCTP is its support for multihomed nodes, i.e., nodes that can be reached under several IP addresses (multihoming). Multihoming allows two endpoints to set up an association with multiple IP addresses for each endpoint. This built-in support for multi-homed endpoints can utilize the redundancy in network, and allow high-availability applications to perform switchover to an alternate path without interrupting the data transfer during link failure situations [1].

If a client is multi-homed, it informs the server about all its IP addresses with the INIT chunk's address parameters. An extension to the SCTP called mSCTP (Mobile SCTP) also

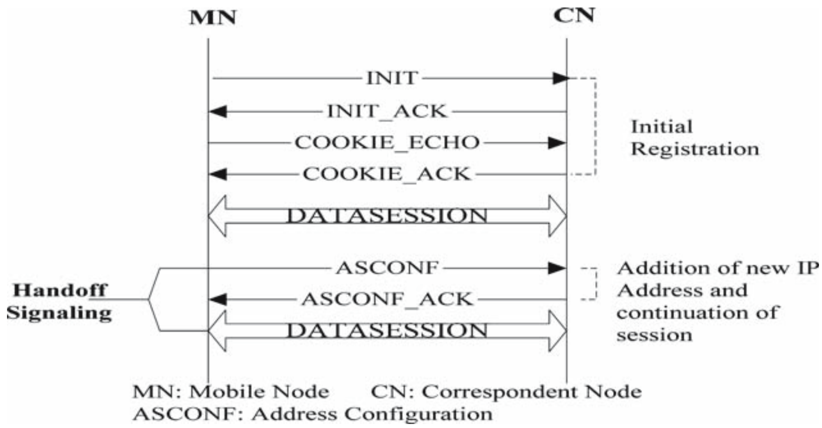

Figure 3. SCTP-based mobility management.

allows dynamic addition and deletion of IP addresses from an association, even if these addresses were not present during association startup. This feature of SCTP is used to support mobility of hosts across different networks.

4. Performance Evaluation of SIP, Mobile IP, and SCTP

4.1. Experimental Network Testbed

We conducted experimental measurements to determine the handoff delay experienced while roaming across different wired/wireless networks. The handoff tests were conducted for each of the mobility protocols: SIP, Mobile IP, and SCTP.

Figure 4 shows the experimental testbed we used to conduct the handoff measurements. The setup consists of a DELL laptop (client machine) equipped with three Network Interface Cards (NICs): a built-in Natsemi Ethernet NIC (100 Mbps), a built-in Orinoco WLAN NIC (11 Mbps) and an external PCMCIA GPRS Sierra Wireless aircard 750 (144 Kbps). The Ethernet interface (`eth0`) of the client machine is connected to a 100 Mbits/sec switch that connects to the external IP network (Internet). The WLAN interface (`wlan0`) of the client machine is associated with a WLAN access point, which is in turn connected to the router for Internet access. The GPRS interface (`ppp0`) is associated with a T-Mobile GPRS base station, which connects to the Internet via the GPRS core network. In order to use the GPRS network, we purchased a GPRS data plan subscription from the T-Mobile service provider [30]. Other components of the testbed include a SIP Registrar server, a HA, and a FA.

The client (MN) and the server (CN) machines were loaded with Redhat [22] 9.0 Linux operating system and used a kernel version of 2.4.20-8. For SCTP-based mobility tests, a user-level implementation of SCTP called Sctplib-1.3.1 [28] was used. For Mobile-IP-based tests, a Mobile IP user-level implementation called Dynamics [19] from Helsinki University of Technology was used. SIP-based mobility was tested by implementing a simple SIP user-agent client [25], a SIP user agent server and a SIP Registrar server using the SIP methods (INVITE, ACK, BYE, REGISTER, and CANCEL) described in RFC 3261 [23].

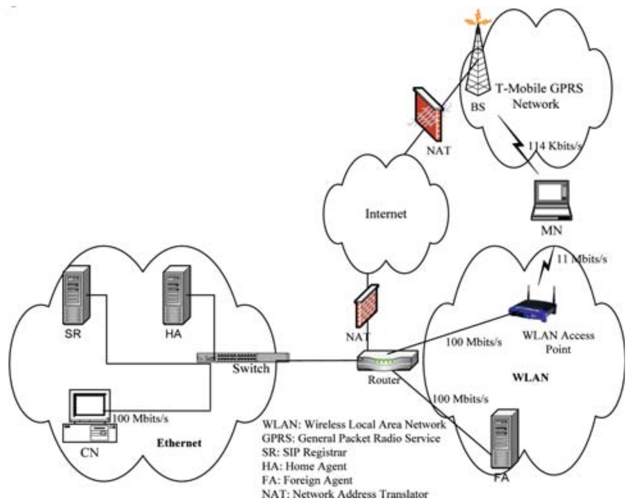

Figure 4. Experimental Testbed used for SIP, Mobile IP, and SCTP handoff performance measurements.

4.2. Measurement Procedures and Performance Metrics

We measured the handoff delay experienced when roaming across three types of networks: Ethernet, WLAN and GPRS by implementing mobility protocols at the application (SIP), network (Mobile IP) and Transport (SCTP) layers.

In the case of SIP, we measured the handoff delay experienced by a Mobile Node in six different cases:

- GPRS to WLAN
- WLAN to GPRS
- Ethernet to WLAN
- WLAN to Ethernet
- Ethernet to GPRS
- GPRS to Ethernet

In the case of SCTP and Mobile IP, we measured the handoff delay in two different cases:

- Ethernet to WLAN
- WLAN to Ethernet

We used shell scripts to automatically deactivate one network interface and activate the other interface. The three network interfaces used were Ethernet (`eth0`), WLAN (`wlan0`) and GPRS (`ppp0`).

The performance metrics that we measured are as follows:

- ***Total Handoff Delay*:** The total handoff delay is the time difference between the last data packet received at the old network interface and the first data packet received on the new network interface. The total handoff delay includes the handoff time as well as the time taken for the first data packet to arrive from the CN to the MN after the MN has switched to the new network.

- ***Handoff Signaling Time*:** The handoff signaling time is a measure of the time required to exchange signaling messages to execute a handoff. The number of signaling messages exchanged is different for each mobility management protocol.
- ***Packet Transmission Delay after handoff*:** The packet transmission delay after the handoff is a measure of the transmission time of a packet from the CN to the MN after the MN has moved to a new network.

All the above measurement tests were conducted while the mobile node was downloading a file from the corresponding node.

The basic handoff implementation, using each protocol, involves the following steps:

- Step 1: Handoff decision: We used the unavailability of the current network to initiate a handoff to the new network.
- Step 2: Acquisition of IP address for the new network interface.
- Step 3: Exchange of handoff messages between various network entities.
- Step 4: Packet transmission/reception following handoff using the new network interface.

For the measurement tests performed, the *Total Handoff Delay* corresponds to the time consumed for the entire handoff process (steps 2 through 4), and the *Handoff Signaling Time* corresponds to step 3 of the handoff process. The *Packet Transmission Delay* following Handoff corresponds to step 4 of the handoff process.

For all tests conducted, we were primarily interested in measuring and comparing the Handoff Delays contributed by each mobility protocol (SIP, Mobile IP, and SCTP). Therefore, authentication was not implemented in our experiments. Moreover, different types of networks use different authentication techniques. For example, GPRS networks employ an International Mobile Subscriber Identity (IMSI) based authentication, while WLANs use various authentication protocols such as Extensible Authentication Protocol (EAP) [1], Protected Extensible Authentication Protocol (PEAP) [2], Lightweight Extensible Authentication Protocol (LEAP) [6], and others. Each of these authentication protocols contributes variable latencies. Hence, we have focused only on handoff performance results of the various mobility protocols without taking into account delays that are associated with authentication/re-authentication procedures during handoffs.

4.2.1. SCTP and mobile IP issues for NAT traversal

It was not possible to measure the handoff delay (for SCTP and Mobile IP) while moving from the GPRS network to the other networks (Ethernet and WLAN) and vice versa because the GPRS operator assigns a dynamic, private IP address to the Mobile Node. A dynamic IP address is one that is not manually specified and is not a permanent address. It is a temporary address that is dynamically configured using the Dynamic Host Configuration Protocol (DHCP). A private IP address is one that can be used by any machine and is therefore re-usable. However, private IP addresses are not routable over the public Internet. They are used in private networks due to the shortage of public, routable IP addresses. The range of IP addresses reserved for private use includes 10.0.0.0 – 10.255.255.255, 172.16.0.0 – 172.31.255.255, 192.168.0.0 – 192.168.255.255. Also, each Internet provider network employs a Network Address Translator (NAT) for providing Internet access to the internal nodes with private IPs and also for security purposes.

The problem with Dynamics implementation of Mobile IP is that it is not "NAT traversal" capable. When a Mobile Node moves to the GPRS network, it acquires a care-of-address

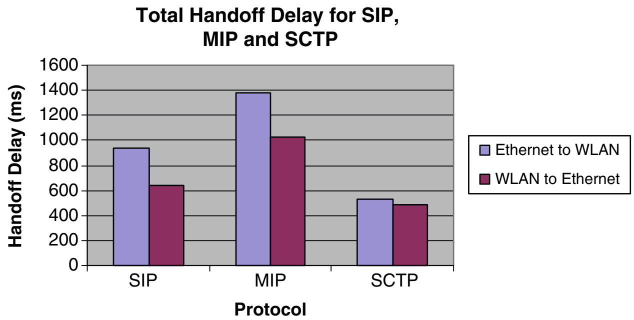

Figure 5. Total handoff delay for SIP, Mobile IP and SCTP.

(CoA), which is a private address. Then the Mobile Node sends a Registration Request to the HA to register its new CoA. However, at the NAT gateway, the private IP address of this packet (source IP address in the IP header) is replaced by the public IP address of the NAT gateway. When the Registration Request arrives at the HA, the HA detects that that the source address of the packet (which is the public address) is different from the CoA inside the Registration Request message (present in the Mobile IP header). As a result, the HA drops the request. Thus, in the case of the Dynamics, it is necessary to have a public, static IP address for the MN. Hence, handoffs involving the GPRS network could not be tested due to the assignment of a private IP.

In the case of SCTP, when the MN is located in the GPRS network and the CN is located on a different network, all packets from the MN have to pass through the NAT. SCTP has certain issues related to NATs. If Network Address Port Translation is used with a multihomed SCTP endpoint, then any port translation must be applied on a per-association basis such that an SCTP endpoint continues to receive the same port number for all messages within a given association. The NAT needs to understand this requirement to allow mobility support using SCTP. Since existing NATs are not designed to support SCTP, a NAT assigns a different port number when the SCTP association changes its primary address. The SCTP server does not accept the change in the port number and breaks the association. Thus SCTP cannot be experimented with a GPRS network employing a NAT that is not configured to support SCTP.

5. Analysis of Experimental Results

In this section we present an analysis of the handoff performance obtained for the three mobility protocols. Figure 5 shows the total handoff delay obtained while roaming from Ethernet to WLAN and vice versa using SIP, Mobile IP, and SCTP. It is worthwhile mentioning that SIP, Mobile IP and SCTP operate at the application, network, and transport layers, respectively.

It can be observed from Figure 5 that the total handoff delay in either direction (Ethernet to WLAN and vice versa) is the lowest in the case of SCTP followed by SIP and is the highest in the case of Mobile IP. The total handoff delay is lowest for SCTP (31% lower compared to SIP and 55% lower compared to Mobile IP for WLAN to Ethernet handoff). The reason for the low handoff delay in the case of SCTP is because the SCTP client immediately adds the IP address of a newly discovered network to its list of available networks and also relays this information to the SCTP server. When a handoff is initiated due to the unavailability of the

Table 1. Components of Handoff Signaling: SIP, Mobile IP, and SCTP.

Protocol	*Handoff Messages*
SIP	Re-Register ACK
Mobile IP	Registration Request, Registration Reply
SCTP	ASCONF_DELETE IP

current network, the client sends an ASCONF_DELETEIP message to the server (to remove the old IP address) and starts using the interface with the new IP address for data transmission. Thus, the handoff process with SCTP involves very few signaling messages thereby resulting in a low total handoff time. Table 1 lists the signaling messages exchanged for implementing handoffs using SIP, Mobile IP and SCTP.

In the case of SIP, when a handoff is initiated, the SIP client sends a Re-INVITE message to the SIP server using the new interface. After the SIP server acknowledges the Re-INVITE, the communication between the client and the server is continued. Thus, handoff delay in the case of SIP is the two-way delay involved in sending the Re-INVITE message and receiving an acknowledgement. We determine the handoff delay at the MN as the time difference between the last data packet received at the old network interface and the first data packet received at the new network interface. Thus, the handoff delay also includes the transmission time of the first packet following the handoff signaling. In the case of Mobile IP, the handoff involves a higher number of signaling messages compared to SIP and SCTP. Mobile IP requires the MN needs to send a Registration Request to the Foreign Agent that forwards the request to the HA. The Registration Reply is sent by the HA to the FA which then gets forwarded to the MN. Due to the high signaling overhead involved in the case of handoffs based on Mobile IP, the signaling time is also higher.

Figure 6 shows the handoff signaling time in the case of the SIP protocol when the MN moves across various networks. It can be observed that the signaling time is the highest when the MN makes a handoff to a GPRS network. The signaling time is comparatively lower when the mobile moves to the WLAN and is the lowest in the case of transition to an Ethernet network. We note that the low signaling delay associated with transition to an Ethernet network is probably because of Ethernet's lowest transmission latency. To confirm this explanation, we performed a simple test using Netperf [17] to determine the available bandwidth and the latency offered by each of these networks. As shown in Table 2, the latency incurred on the GPRS network is comparatively higher as compared to Ethernet and WLAN. This accounts for the high-handoff signaling delay when the MN moves to the GPRS network. We also observe (from Figure 6) that there is a 41% reduction in the handoff signaling time in the case of SIP when compared to Mobile IP (for handoff to a WLAN) and a 60% decrease in the handoff signaling time in the case of SCTP as compared to Mobile IP.

Figure 7 shows the transmission delay incurred by packets arriving at the MN after the handoff. We observe that in the case of Mobile IP, we obtained highest packet transmission delay. As observed from Figure 7, there is a 47% decrease in the packet transmission delay in the case of SIP as compared to Mobile IP (in the case of handoff to a WLAN) and a 54% decrease in the packet transmission delay with SCTP as compared to Mobile IP (in the case of handoff to a WLAN). This is because, after handoff, packets from the CN to the MN have to be routed through the HA and the FA before they can reach the MN. This introduces additional

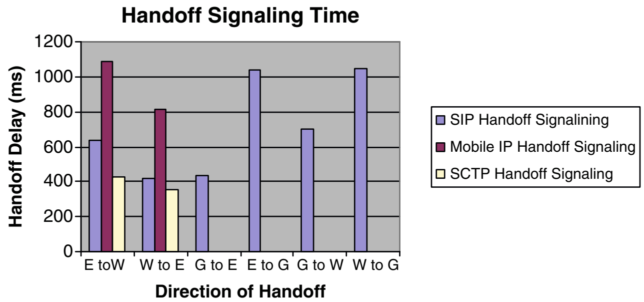

Figure 6. Handoff signaling time for SIP, Mobile IP, and SCTP.

Table 2. Network characteristics determined using Netperf

Network Type	*Link Speed*	*Actual Measured Bandwidth*	*Average Latency (one-way)*
GPRS	114 Kbits/s	28.9 Kbits/s	891 milliseconds
WLAN	11 Mbits/s	5.51 Mbits/s	61 milliseconds
Ethernet	100 Mbits/s	88.8 Mbits/s	36 milliseconds

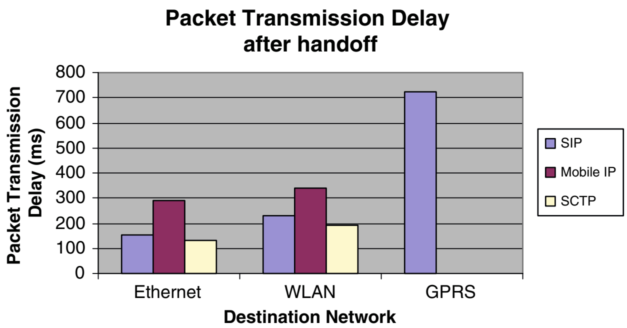

Figure 7. Packet transmission delay after handoff.

delay in the transmission time. The packet transmission delay for SCTP and SIP is almost the same. In both of these cases, packets following handoff are sent directly from the CN to the MN. This results in a lower packet transmission delay for SIP and SCTP as compared to Mobile IP. It is worthwhile noting that the Mobile IP issue on routing packets through the HA and the FA is solved by *Route Optimization*, which involves sending binding updates to inform the CN of the actual location of the MN. However, route optimization has other problems [31] associated with it and has not been widely deployed in existing nodes and therefore the binding update function is generally not supported. This is the reason why we have only considered the base Mobile IP protocol implementation for our experimental evaluations.

6. Conclusion

In this paper we empirically compared the handoff performance of three of the most popular mobility protocols: SIP, Mobile IP and SCTP. We found that SCTP performs well both, in terms of handoff delay, as well as the packet transmission time after a handoff. The SIP protocol incurred a higher handoff delay compared to SCTP but the packet transmission time for packets after a handoff was almost comparable for the two protocols. Mobile IP incurred higher handoff delay as well as longer packet transmission time following handoff to a new network. However, Mobile IP keeps the change in the IP address completely transparent to the other end-system. In the case of SIP and SCTP, the change in the destination IP address has to be conveyed to the node at the other end. SCTP-based mobility is however completely transparent to the application, but with SIP, applications need to be aware of mobility. Furthermore, SCTP can be used in scenarios where a mobile client initiates a session with a fixed server. In order to support peer-to-peer services (for example involving mn at both communicating ends) SCTP needs to be used with additional location management schemes.

We also discussed some issues related to the deployment of Mobile IP and SCTP over networks using private IP addresses and NATs. Mobile IP and SCTP are not capable of operating in networks with NAT mechanisms. Since many network operators use NATs in their networks, it is crucial to extend these protocols to enable them to operate across heterogeneous domains. One method that can be used to enable this feature is to use UDP encapsulation in each of these protocols. Since most NATs are already designed to provide support for UDP packets, encapsulating SCTP packets inside UDP can make SCTP operate across NATs belonging to different network domains. However, this would introduce additional encapsulation-decapsulation delays.

7. Future Work

Our future research efforts will focus on the design of a solution for handoff management, that is, not specific to a single layer of the network protocol stack, but exploits a cross-layer design to achieve seamless handoffs across heterogeneous networks. We plan to implement a mobility middleware that performs handoffs using the information from various layers (such as link quality information from layer 2, QoS information from layer 4, etc.) of the protocol stack and deployable across existing network infrastructures.

Other recent works that have investigated techniques aimed at reducing handoff latency include Seamoby [7], Hierarchical Mobile IPv6 (HMIPv6) [27], and Host Identity Protocol (HIP) [15]. Recently, the IEEE 801.21 (Media Independent Handover) [16] Working Group has also undertaken the task of addressing mobility issues related to performing seamless handoffs across heterogeneous access technologies such as 802.11, 802.16, 802.3, and cellular networks.

Acknowledgments

This work was supported by grants from Sun Microsystems (Palo Alto) (EDUD-7824-000145-US), Microsoft Corporation (Seattle), and Ixia Corporation (Calabasas). We are grateful to the anonymous reviewers for their suggestions and comments which helped to improve the quality and presentation of this paper.

Appendix- List of Acronyms

BNF: Backus-Naur Form; CN: Correspondent Node; CoA: Care-of-Address; DHCP: Dynamic Host Configuration protocol; EAP: Extensible Authentication Protocol; FA: Foreign Agent; GPRS: General Packet Radio Service; GSM: Global System for Mobile Communication; HA: Home Agent; IETF: Internet Engineering Task Force; IMSI: International Mobile Subscriber Identity; IP: Internet Protocol; LAN: Local Area Network; LEAP: Lightweight Extensible Authentication Protocol; MN: Mobile Node; MSCTP: Mobile Stream Control Transmission protocol; NAT: Network Address Translation; NIC: Network Interface Card; PEAP: Protected Extensible Authentication Protocol; SCTP: Stream Control Transmission Protocol; SIP: Session Initiation Protocol; TU: Transaction User; WLAN: Wireless Local Area Network.

References

1. B. Aboba and D. Simon, "PPP EAP TLS Authentication Protocol", RFC 2716, 1999.
2. H. Andersson, S. Josefsson, G. Zorn, and B. Aboba, "Protected Extensible Authentication Protocol", Internet Draft October 2001, http://tools.ietf.org/html/draft-josefsson-pppext-eap-tls-eap-01.txt.
3. "Backus-Naur Form", http://www.cacs.louisiana.edu/∼mgr/404/burks/foldoc/99/9.htm.
4. N. Banerjee, S. Das, and A. Acharya, "SIP-Based Mobility Architecture for Next Generation Wireless Networks", in *Proc. of the third IEEE International Conference on Pervasive Computing and Communications*, pp. 181–190, 2005.
5. F. Chiussi, A. Khotimsky, and S. Krishnan, "Mobility Management in Third-generation all-IP networks", *IEEE Communications Magazine*, Vol. 40, No. 9, pp. 124–135, 2002.
6. Cisco et al., "Cisco LEAP Protocol Description", http://lists.cistron.nl/pipermail/cistron-radius/2001-September/ 002042.html, 2001.
7. Context Transfer, Handoff Candidate Discovery, and Dormant Mode Host Alerting (seamoby), http://www.ietf.org/html.charters/OLD/seamoby-charter.html.
8. W. Eddy, "At What Layer Does Mobility Belong", *IEEE Communications Magazine*, Vol. 42, No. 10, pp. 155–159, 2004.
9. H. Fathi et al., "On SIP Session Setup Delay for VoIP Services Over Correlated Fading Channels", *IEEE Transactions on Vehicular Technology*, Vol. 55, Issue: 1, pp. 286–295, 2006.
10. R. Fracchia et al., "A WiSE Extension of SCTP for Wireless Networks", in *Proc. of IEEE International Conference on Communications*, Vol. 3, pp. 1448–1453, 2005.
11. S. Fu and M. Atiquzzaman, "SCTP: State of the art in research, products, and technical challenges," *IEEE Communications Magazine*, Vol. 42, No. 4, p. 64–76, 2004.
12. J. Funasaka et al., "A Study on the Primary Path Switching Strategy of SCTP", in *Proc. of Autonomous Decentralized Systems*, pp. 536–554, Beijing, China, 2005.
13. S. Gokhale and J. Lu, "Signaling Performance of SIP-based VoIP: A Measurement-based Approach", in *Proc. of the Global Telecommunications Conference*, Vol. 2, pp. 761–765, Tampa, Florida, USA, 2005.
14. E. Hernandez and A. Helal, "Examining Mobile IP Performance in Rapidly Mobile Environments: The Case of a Commuter Train", in *Proc. of 26th IEEE Local Computer Networks (LCN'01)*, FL, 2001, pp. 365–372.
15. "Host Identity Protocol", http://www.ietf.org/html.charters/hip-charter.html.
16. "IEEE 802.21 Template", http://www.ieee802.org/21/.
17. R. Jones, Netperf, http://www.netperf.org.
18. S. Liu et al., "Collaborative SCTP: A Collaborative Approach to Improve the Performance of SCTP over Wired-cum-Wireless Networks", in *Proc. of the 29th Annual IEEE International Conference on Local Computer Networks*, pp. 276–283, Providence, Rhode Island, USA, 2004.
19. Mobile IP, Dynamics Implementation, http://www.cs.hut.fi/Research/Dynamics.
20. T. Park and A. Dadej, "OPNET simulation modeling and analysis of enhanced Mobile IP", in *Proc. of Wireless Communications and Networking Conference*, Vol. 2, pp. 1017–1024, Providence, Rhode Island, USA, 2003.
21. C. Perkins, "Mobile Networking through Mobile IP", *IEEE Internet Computing,* Vol. 2, No. 1, pp. 58–69, 1998.

22. "Red Hat", http://www.redhat.com.
23. J. Rosenberg et al., "Session Initiation Protocol", RFC 3261.
24. A. Saleh, "Mobile IP performance and Interworking architecture in 802.11 WLAN/CDMA2000 Networks", in *Proc. of Second Annual Conference on Communication Networks and Services Research*, 2004, pp. 75–79.
25. H. Schulzrinne, "Application Layer Mobility with SIP", *ACM SIGMOBILE Mobile Computing and Communications*, Vol. 4, No. 3, pp. 47–57, 2000.
26. J. Shi et al., "Experimental Performance of SCTP in Wireless Access Networks", in *Proc. of International Conference on Communication Technology*, Vol. 1, pp. 392–195, 2003.
27. H. Soliman, K. Malki, and L. Bellier, "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)", RFC 4140, 2005.
28. R. Stewart and Q. Xie et al., "Stream Control Transmission Protocol", RFC 2960, 2000. SCTP Implementation, http://www.sctp.de.
29. R. Stewart et al., "Stream Control Transmission Protocol", RFC 2960.
30. T-Mobile, http://www.tmobile.com.
31. E. Wedlund, "Mobility Support Using SIP", http://www.cs.columbia.edu/IRT/papers/Wedl9908_Mobility.pdf.
32. W. Wu, N. Banerjee, K. Basu, S. Das, "SIP-based Vertical Handoff between WWANs and WLANs", *IEEE Wireless Communications*, Vol. 12, pp. 66–72, 2005.
33. C. Yeh et al., "SIP Terminal Mobility for both IPv4 and IPv6", in *Proc. of the 26th IEEE International Conference on Distributed Computing Systems and Workshops*, pp. 53–59, 2006.
34. L. Zhang, J. Cao, and K. Das, "A Mailbox-based scheme for Improving Mobile IP Performance", In *Proc. of 23rd International Conference on Distributed Computing Systems Workshops*, pp. 864–869, 2003.

Sherali Zeadally received the BA and MA degrees in Computer Science from University of Cambridge, England, and his doctoral degree in Computer Science from University of Buckingham, England, in 1996. He is currently an Associate Professor in the Department of Computer Science and Information Technology at the University of the District of Columbia, Washington DC where he directs the Network Systems Laboratory. He currently serves on the editorial boards of several international journals and he is the Co-Editor-in-Chief of the International Journal of Internet Protocol Technology. He is a Fellow of the British Computer Society (FBCS) and a Fellow of the Institution of Electrical Engineers (FIEE), UK. His research interests include Computer Networks (wired, wireless), Mobile Computing, Network and System Security, Ubiquitous Computing, Multimedia Networking, and Performance Evaluation of Networks and Systems.

Farhan Siddiqui received a MS degree in Computer Science from Wayne State University, Detroit, Michigan, and a BE degree in Computer Science from Osmania University, India. She is currently completing her PhD in wireless networks in the Department of Computer Science at Wayne State University. Her research interests include Wireless Networks, Mobile Computing, Voice over IP, and Network Security.