ADAPTIVE END-TO-END MOBILITY SCHEME FOR SEAMLESS HORIZONTAL AND VERTICAL HANDOFFS

Abdellatif Ezzouhairi, Alejandro Quintero, Samuel Pierre
Mobile Computing and Networking Research Laboratory (LARIM)
Department of Computer Engineering, École Polytechnique de Montréal
P.O. Box 6079, succ. Centre-Ville, Montreal, Quebec, H3C 3A7, Canada
Phone: (514) 340-3240 ext. 4685. Fax: (514) 340-3240
E-mail: {Abdellatif.Ezzouhairi; Alejandro.Quintero; Samuel.Pierre}@polymtl.ca

ABSTRACT

Mobility management constitutes one of the most significant task to be investigated for Next Generation Mobile Networks (4G). Motivated by connectivity facilities and flow control offered at the transport layer, a number of Stream Control Transmission Protocols (SCTPs) based mobility schemes have been proposed to handle this important issue. However, these proposals are hindered by drawbacks such as unnecessary handoff delays incured by horizontal handoffs. Moreover, the throughput measured immediately after a handoff is affected quite considerably by spurious retransmissions due to failed Selective Acknowledgment messages (SACKs) and data retransmission lost. This paper proposes a new Hierarchical Transport layer Mobility protocol (HTM) that deals with local and global mobility and improves throughputs during the handoff period. HTM exploits the dynamic address reconfiguration feature of SCTP and introduces an Anchor Mobility Unit (AMU) in order to complete more efficient handoff procedures. Simulation and numerical results reveal that HTM guarantees lower handoff latency and packet loss, good throughput and limited signaling load compared to mSCTP (mobile SCTP) based mobility.

Keywords: Heterogeneous networks, mobility management, SCTP, end-to-end roaming.

1 INTRODUCTION

The next generation of mobile communication systems, referred to as 4G, 3G+ or beyond 3G, is intended to integrate both current and emerging mobile networks around an IP backbone. For example, this will include second and third generation cellular networks (2G and 3G), satellite systems, Wireless Local Area Networks (WLANs), amongst others. Since each technology is tailored to reach a particular market or a specific type of user services, integrating these heterogeneous systems becomes highly interesting as they offer many possibilities to increase bandwidth, Internet accessibility and area coverage. For example, a mobile user may choose to access a WLAN to send a large data file, but selects a 3G cellular network to place a voice call. However, implementing this type of integrated system implies numerous challenges in mobile handset design, wireless system discovery, terminal mobility, security and billing [1]. Mobility management remains the most significant task to be investigated since it aims to guarantee mobile users disruption-free connections while roaming through heterogeneous networks. Traditionally, mobility management comprises location management and handoff management [4].

Location management is a process which allows networks to localize mobile users' current attachment point for data delivery.

Handover or handoff management enables the network to sustain mobile user connections, while they move and change network access points.

Handoff mechanisms are usually categorized into: hard and soft handoffs. A hard handoff, also known as break-before-make, is completed by first disconnecting with the current access point before switching to another one. This type of handoff mechanism is particularly suitable for delay-tolerant communications traffic. On the other hand, the soft handoff also known as make-before-break, is employed by establishing a connection with a new access point before disconnecting from the existing point of attachment. This category of handoff mechanism is particularly suitable for handling latency-sensitive communication services such as videoconferencing. In this sense, Mobile IP [6] and its further enhancements such as HMIPv6 [7], FMIPv6 [8] and FHMIPv6 [9] are considered among the IETF standards widely accepted to deal with mobility management. However, this category of mobility schemes suffers from weaknesses such

as handoff latency, packet loss and signaling load pertaining to the number of bindings to be executed. In addition, certain mobility schemes based on TCP [10] and SIP [11] have been investigated as alternate solutions to the traditional mobile IP. Generally, these proposals need tremendous modifications in both protocol stacks and network architecture [12]. With the standardization of SCTP [13], and more particularly with its novel ADDIP Extensions [14], more attention has been paid to experiment mobility over the transport layer. Actually, the transport layer mobility schemes do not depend on the underlying infrastructures and offers the possibility to control the flow and to pause transmission in expectation of a handoff. Thus, a number of solutions which exploit the multihoming features of SCTP have been introduced. Yet, to the best of our knowledge, none of these proposed approaches deal with local mobility at the transport level. This means that current SCTP-based mobility proposals focus on the multihoming feature and do not consider the fact that most of the MN's handoffs are completed inside the same wireless technology (i,e., horizontal handoff). Note that inside an homogeneous technology, an MN may not simultaneously use its two wireless interfaces for communication [15]. Obviously, this leads to superfluous delays due to L2 handoff, movement detection, authentication and address configuration. Moreover, certain hidden effects pertaining to fast handovers, such as failed SACKs (Selective Acknowledgements) are not addressed.

The main concern of this paper is to propose a new Hierarchical Transport layer Mobility scheme (HTM) that takes into account local and global mobility in order to reduce handoff latency, packet loss and signaling costs. Additionally, the problem of spurious retransmissions due to failed SACKs and data retransmission lost is addressed. Finally, several simulations and an analytical model are investigated in order to demonstrate the effectiveness of the proposed mobility scheme. In the rest of this paper, the terms mobile user and mobile node will be used interchangeably.

The remainder of this paper is structured as follows: Section 2 presents related work and Section 3 describes the proposed mobility scheme. An analytical model is introduced in section 4. Performance analyses and simulation results are presented in Section 5. Finally, Section 6 concludes the paper.

2 RELATED WORK

The IP layer is traditionally considered as the default place where mobility is implemented since the IP protocol remains widely used to connect heterogeneous communication systems. However, an increasing interest is recently given to experience mobility at the transport and application levels. In this section, we give an overview of the well-known mobility mechanisms available in the literature.

2.1 IP layer mobility

Traditionally, mobility management is performed at the network layer due to the use of the Internet Protocol (IP) that allows routing packets between different technologies. In this context, several approaches propose coping strategies for IP layer mobility. Among these, Mobile IPv6 (MIPv6) is the most popular mechanism that allows mobile nodes to remain reachable in spite of their movements within IP-based mobile environments. However, MIPv6 has some well-known drawbacks, such as high signaling overhead, packet loss and handoff latency, thereby causing real-time traffic deterioration which can be perceived by users [17]. These weaknesses led to the investigation of other solutions designed to enhance MIPv6. The IETF proposed new MIPv6 extensions including Hawaii [18], Cellular IP [19] and Hierarchical MIPv6 (HMIPv6). These protocols tackle intra-domain or micro-mobility, while MIPv6 is used for inter-domain or macro-mobility. However, this solution generates extensive bidirectional tunneling as long as the mobile moves inside the same administrative domain. Additionally, FMIPv6 was proposed to reduce handoff latency and minimize service disruption during handoffs pertaining to MIPv6 operations, such as movement detections, binding updates and address configurations. Although FMIPv6 paves the way for improving MIPv6 performance in terms of handoff latency, it does not efficiently reduce signaling overhead (due to new messages being introduced and exchanged for handoff anticipation) nor does it prevent packet loss (due to space requirements). This may lead to unacceptable service disruptions for real time applications. Combining HMIPv6 and FMIPv6 motivates the design of Fast Handover for HMIPv6 (FHMIPv6) to increase network bandwidth efficiency. However, FHMIPv6 may inherit drawbacks from both HMIPv6 and FMIPv6, those pertaining to synchronization and signaling overhead issues, for instance. Furthermore, the IETF has also proposed a network-based mobility referred to as Proxy Mobile IPv6 [5] to ensure mobile user roaming without its participation in any mobility-related signaling. However, this type of mobility schemes depends entirely on the network infrastructure and need a permanent bidirectional tunnel between the MN and CN.

2.2 Application layer mobility

Handling mobility at the application layer has also received a lot of attention since this category of solutions is almost independent of the underlying technologies. To accomplish this type of mobility,

the SIP [4] protocol is widely used. Thus, when a mobile node moves during an active session into different network, it first receives a new address, and then sends a new session invitation to its correspondent node. Subsequent data packets are forwarded to the MN using this new address. However, SIP by itself does not guarantee the maintenance of established Transmission Control Protocol (TCP) sessions or User Datagram Protocol (UDP) port bindings when moving, so further extensions such as S-SIP [20] are needed to provide seamless handover capabilities.

2.3 Transport layer mobility

Recently, transport layer-based mobility is gaining attention since it does not require a concept of home network and mobile nodes can perform smooth handovers if they are equipped with multiple interfaces. Moreover, this category of mobility schemes may benefit from flow control and the possibility to pause transmission during the handoff period. The first transport layer mobility solutions were based on TCP, and then other interesting mobility approaches have been proposed with the standardization of SCTP [13] and mSCTP [14].

2.3.1 TCP-based mobility

In the last few years, several transport layer mobility schemes have been proposed to benefit from the connectivity facilities and flow control offered at the transport level. From this perspective, a new TCP protocol architecture was proposed to support mobility [22]. However, tremendous changes must be performed over the entire network to reach this goal. MSOCKS [23] is another TCP-based proposal which does not require changes to the network layer infrastructure. However, it suffers from high latency and packet loss, since it follows a make-after-break approach (disable MN connections until a new path is ready). Migrate [10] is another TCP-based mobility solution which aims to ensure transparent TCP connection migration. Nevertheless, this solution requires changes to TCP implementation at both ends of the connection. Multi-homed TCP, introduced by [24], aims to use several addresses in parallel for the same connection by proposing to use new TCP Protocol Control Bloc (PCB) to name the TCP socket, thereby allowing underlying IP addresses to change. However, this approach needs huge modifications and remains, accordingly, not used.

2.3.2 SCTP-based mobility

Performing mobility on the transport layer becomes more realistic with the emergence of the Stream Control Transmission Protocol (SCTP), and even more so with its mobile extension. Indeed, SCTP is a new transport layer protocol that was recently standardized under the RFC 4960. It inherited many TCP properties, but it also introduces novel and interesting features, such as

multistreaming and multihoming. Multistreaming consists of delivering independent data streams by decoupling reliable deliveries from message ordering. This feature prevents receiver head-of-line blocking in cases where multiple independent data streams occur during a single SCTP session. On the other hand, multihoming allows an SCTP node to be reached through multiple IP addresses (interfaces). In fact, two SCTP nodes can exchange data by defining a common association. In SCTP terminology, an association is equivalent to a TCP connection. End points can be single-homed or multihomed. When single-homed, SCTP nodes are defined as [IP address: SCTP port], otherwise they are designated as [IP1 address, IP2 address…IPn address: SCTP port]. When establishing an association, end points define their primary path, as well as the secondary ones. The primary path is used to transfer data, while secondary paths are used for retransmissions and backups in the event of primary path failures. The SCTP ADDIP [14] Extension enables SCTP nodes to dynamically add, delete and modify their primary address without terminating an ongoing association.

In [26] the authors propose an approach to ensure vertical handoffs between UMTS and WLAN networks using SCTP multi-homing capabilities. In [27], a TraSH mobility scheme was proposed to perform seamless handovers between heterogeneous networks. In SIGMA [28], the authors propose an SCTP-based mobility architecture that integrates location management to ensure seamless handovers. In [29], the authors advance certain triggering rules to improve throughput during SCTP-based handoffs. All of these proposals are based on the mobile SCTP extension (mSCTP) and their corresponding mobility procedure is summarized in Fig. 1.

> 1. *Obtain an IP address from a new location.*
> 2. *Add the new IP address to the association.*
> 3. *Change the primary IP address.*
> 4. *Delete the previous IP address from the SCTP association.*

Figure 1: Mobile SCTP-based handoff procedure

In [30][31], the authors put forward new transmission techniques by attempting to enable SCTP-based mobility schemes with concurrent multi-path data transfers. Unfortunately, all of the proposed schemes focus on the inter-system handoffs (i,e., vertical handovers) and do not consider the fact that the majority of handoffs are performed inside the same wireless system (i,e., horizontal handoffs). Accordingly, mobile users must endure unnecessary handoff delays and signaling loads which may become significant in case of frequent handovers. Moreover, a number of hidden effects such as spurious retransmissions due to failed SACKs and data lost reduce considerably

throughput during handoff periods.

Besides the aforementioned proposals, the Host Identity Protocol (HIP) is introduced to operate in a new layer between the network and the transport layers. The HIP protocol aims to separate the identity (end points and host identifiers) and location information (IP routing) by introducing a new name-space, the Host Identity (HI). The HI is basically a public cryptographic key of a public-private key-pair. A host possessing the corresponding private key can prove the ownership of the public key, i.e. its identity. The separation of the identity and locator makes it is also simpler and more secure to handle mobility and multi-homing in a host. However, this kind of solution suffers from high overhead for short transactions (handshake) and lack of micro-mobility.

3 HIERARCHICAL TRANSPORT LAYER MOBILITY (HTM)

This section offers a detailed description of the proposed Hierarchical Transport layer Mobility (HTM) that copes with local and global mobility at the transport level and addresses the problem of deterioration of throughput during the handoff period. More specifically, a functional scenario is first introduced. Then, the various elements pertaining to the proposed HTM are presented. Note that security issue is out of the scope of this paper.

3.1 Functional Scenario

This subsection presents a functional scenario that aims to outline some critical issues that must be addressed when designing a novel SCTP-based mobility scheme.
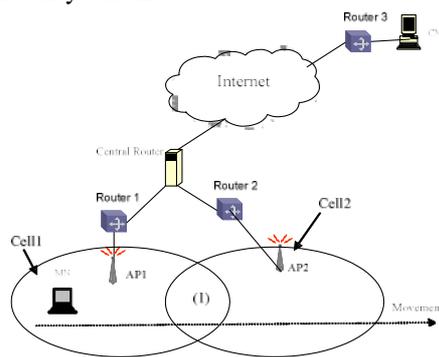

Figure 2: Functional scenario

Fig. 2 illustrates a very common scenario for an MN roaming through homogeneous networks. We assume that the MN is multihomed and equipped with two wireless interfaces. The MN and CN are supposed to support the SCTP protocol.

Initially, the MN has established an association with CN and exchanges its data through AP1. Once the MN enters into the overlapping area (Position

(1)), it initiates a horizontal handoff (intra-system) based, for instance, on the quality of the received signal. However, in most radio systems, the MN cannot simultaneously use its two interfaces when it moves inside a same wireless technology. Hence, handoff latency, in this case, will include delays relevant to L2 link switching, movement detection, address configuration and association updates. Thus, without taking into account local handoffs, the MN incurs unnecessary handoff delays. Moreover, when an MN changes its primary path, a number of SACKs sent to the MN's previous location are lost as it is shown in Fig. 3. Note that the same situation occurs when the CN acts as the receiver.
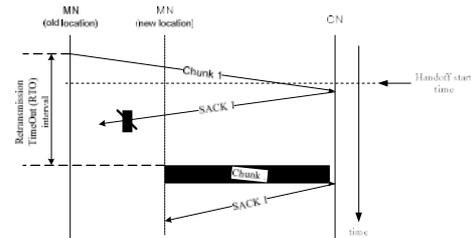

Figure 3: Example of failed SACK due to primary path changes

Indeed, the RFC 4960 states that "an endpoint SHOULD transmit reply chunks (e.g., SACK, HEARBEAT ACK, etc.) to the same destination transport address from which it received the DATA or control chunk to which it is replying; and when its pair is multihomed, the SCTP endpoint SHOULD always try to send the SACK to the same destination address from which the last DATA chunk was received". As a result, a number of SACKs transmitted through a previous path fails to reach their destination since the MN has changed its primary IP address. Consequently, unnecessary Congestion Window (CWND) reductions ensue. Under such circumstances, one may expect that the throughput will be affected. Additionally, when the MN operates as a receiver, a number of data chunks sent to the MN's old primary path will be lost due to a handoff event. Furthermore, all data retransmissions (chunks) performed after the expiration of the retransmission timeout (RTO) will be also lost as it is shown in Fig. 4. Accordingly, a reduction of the CWND parameter will follow. It is clear that such a phenomenon will have a serious impact on the throughput observed during the handoff period.
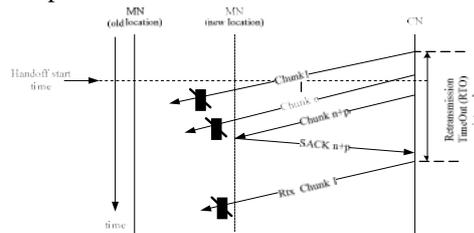

Figure 4: Example of failed chunks due to primary path changes

The following section introduces our proposed hierarchical mobility mechanism that deals with local and global roaming, and addresses the problem of spurious retransmissions due to failed SACKs and data chunks.

3.2 HTM Architecture

In order to address the aforementioned drawbacks, we propose a novel Hierarchical Transport layer Mobility scheme (HTM) that considers local and global mobility. More specifically, HTM aims to exploit existing hierarchical topologies to implement its new Anchor Mobility Unit (AMU) which allows mobile users to perform local handoffs. In fact, topologies that use hierarchical routers (as illustrated in Fig. 2) are frequently encountered in wireless network designs. Hence, routers (or central routers) that may integrate AMU functionalities can be easily found. Basically, HTM consists of a two-unit handoff procedure designed as: HTM^{local} and HTM^{global}. The former treats local/intra-domain mobility, while the latter deals with global/inter-domain roaming.
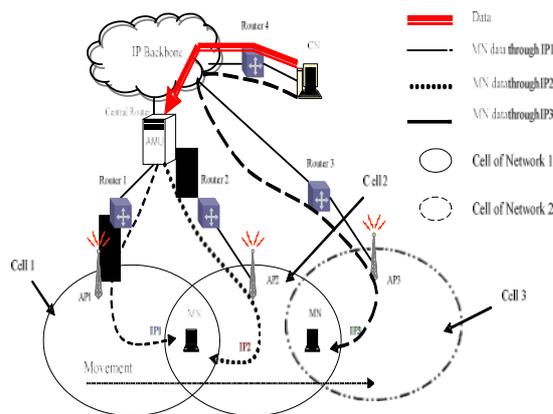

Figure 5: HTM architecture

The HTM architecture that supports both local and global handoffs is illustrated in Fig. 5. In this architecture, the MN is assumed to be multihomed with two active wireless interfaces. Initially, the MN is assigned to Cell 1 and receives data from its Correspondent Node (CN) on its IP1 interface. While moving, the MN changes its point of attachment from Cell 1 to Cell 2 and finally to Cell 3. When the MN hands off from Cell 1 to Cell 2, it performs a local/intra-domain handoff. However, when it moves from Cell 2 to Cell 3, it completes a global/inter-domain handover. Additionally, AP1 and AP2 belong to the same wireless system, while AP3 belongs to an external mobile system. Router1 and Router2 are connected to a Central Router (CR) which supports AMU functionalities. The main role

of the AMU unit consist of assisting mobile nodes to perform seamless handoffs. Each AMU is identified by an AMU-ID (AMU-Identifier), which is periodically broadcasted in the AP/AR beacons. AMU-IDs are highly useful for MNs to decide whether to perform local or global handoff. Basically, the AMU functionalities consist of buffering traffic during the disruption period and performing redirection when the MN is attached to the new link. The AMU process is depicted in Fig. 6.

More specifically, the AMU continuously listens to the redirect events (Redirect-Init). Once a Redirect-Init event occurs, the AMU starts buffering traffic sent to the old MN's IP address. When the MN is attached to its new location, it sends a Redirect-Ready message to notify the AMU that it is ready to receive data on its newly configured IP address. The AMU redirect process ends when no more packets are sent to the old MN address. The following section provides further details pertaining to the proposed handoff procedures when dealing with local and global mobility.
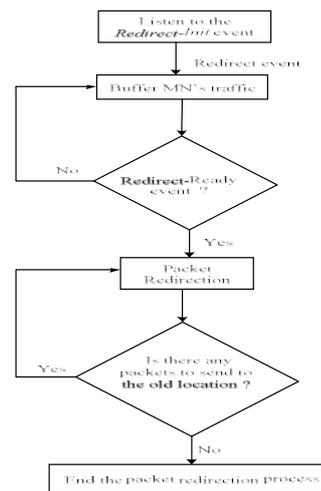

Figure 6: AMU redirection process

3.3 HTM Handoff Procedures

To take benefit of the SCTP multihoming feature we have to remember that when a mobile node moves between cells belonging to a same technology, it can use only one wireless interface a time. However, the MN can simultaneously use its two wireless cards when it moves through cells belonging to heterogeneous technologies. Thus, if we take into account the fact that mobile devices will become increasingly powerful, intelligent and sensitive to link changes, we can assume that the MN detects its movement toward a new access router by using L2 triggers (ie., weak signal strength, high bit error rate, etc.).

As pointed out earlier, the MN detects the presence of the AMU unit through the periodic beacons received from its current point of attachment. Hence, when the MN receives L2 trigger, it sends a RAS_req (Router Address Solicitation request) message to its serving AMU to obtain a new address from the next access router (NAR). Accordingly, if the MN receives a new IP address, it concludes that it has to perform an HTM^{local} procedure (local handoff). Otherwise, it runs the HTM^{global} procedure (global handoff).

3.3.1 HTM Local Handoff Procedure (HTM^{local})

The HTM^{local} procedure is initiated when an MN perform a handoff, for example from Cell 1 to Cell 2, as illustrated in Fig. 5. In this case, it obtains an IP address from AR2 through its serving AMU unit. Practically, this task can be completed with DHCP [32] or IPv6 autoconfiguration [33]. The AMU keeps an association between the new obtained address and the one currently used by the MN. From this moment, the MN is ready to perform a handoff. Recall, that until now the MN continues to receive data from its old path. When the MN decides to move to its new location, it sends a Redirect-Init message to the AMU unit. This message informs the AMU that the MN is performing a L2 link switching (L2 handoff). At this time, the AMU buffers all the packets sent to the MN's previous address until the MN attaches to NAR's link. As soon as the MN is attached to the new access router (NAR), it sends a Redirect-ready message to notify the AMU that it has been successfully attached to its new location. Upon receiving the Redirect-ready message, the AMU starts packet forwarding to the new MN's IP address. At the same time, the MN sends an ADDIP_Soft chunk to inform the CN that a handoff had occurred and it has to set the new MN's IP address as the primary path of their association. Finally, when the MN is completely far from its previous attachment point, the old path is deleted. The entire HTM^{local} procedure is illustrated in Fig. 7.
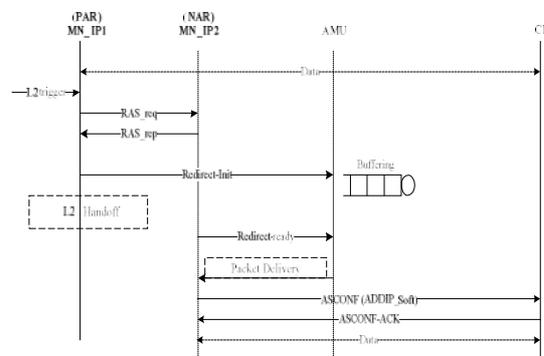

Figure 7: The HTM^{local} handoff procedure

ADDIP_Soft is a new chunk introduced to

perform the set of primary path when the MN is subject to a local handoff. When the CN receives the ADDIP_Soft chunk, it concludes that its pair (MN) has performed a local handoff. The CN immediately transmits packets through the MN's new IP address (IP2) and ignores the previous one (IP1). The description of the new proposed ADDIP_Soft chunk appears in Fig. 8.

Type = 0xC008	Length = 20
Chunk-ID = 0x11122233	
Value = 0x0a010101 (New address)	
Value = 0x0a010111 (Old address)	

Figure 8: Description of the ADDIP_Soft chunk

The main advantage of the proposed HTM^{local} consists of allowing the MN to perform fast handoffs when an AMU component is available. Note that the tunnel established between the AMU and CN operates only during the handoff period. This approach is completely different from HMIPv6 and Proxy Mobile IP principles where the tunnel is maintained as long as the MN moves inside the same administrative domain. Additionally, the Network Address Translation (NAT) concept is not suitable in our case since NAT is not designed for mobile purposes. Moreover, many applications and protocols need to use real end-to-end IP addresses. For instance, this is the case with IP security architecture that cannot work across a NAT device since the original headers are digitally signed. The proposed HTM is expected to reduce latency and limit signaling load over the network. Additionally, the problem of spurious retransmissions due to failed SACKs is considered since all messages (including SACKs) destined to the MN are forwarded to the MN through the AMU unit. Finally, note that the AMU unit is implemented over an existing architecture. Hence, in cases where adding an AMU component would be impossible, the MN can perform its handovers by using the HTM^{global} procedure.

3.3.2 HTM Global Handoff Procedure (HTM^{global})

In the absence of an AMU unit, all handoffs are completed with the HTM^{global} procedure described in Fig. 9. However, handoffs performed in this case (i.e., without an AMU) may be either horizontal (i.e., same technology) or vertical (i.e., different technology). When the handoff is performed within a same technology (i.e., horizontal handover), the handoff disruption time will include: L2 handoff movement detection, authentication, address configuration and association update (i,e., ADDIP and Set-Primary signaling messages). However, when the MN performs a vertical handover, the two wireless interfaces can be used simultaneously. Thus, L2 handoff, movement detection,

authentication, address configuration and association update (ADDIP), can be completed while the MN continues to receive traffic on its old path. When an MN wants to perform a handoff, for example, from Cell 2 to Cell 3 (refer to Fig. 5), it listens to the AP3 beacons. Then, it obtains a new IP address from AR3 (i.e, IP3) to configure its second wireless interface.

The rest of the handoff signaling procedure, in the absence of an AMU unit, is given as follows:

1. The MN sends an ASSCONF (ADD IP) message to inform the CN that to add a new IP (MN IP3) address to their association.
2. The CN responds with an ASSCONF-ACK acknowledgement.
3. The MN asks the CN to consider IP3 as its primary address by sending the ASSCONF (Set Primary Address) chunk.
4. The CN sets the new IP address as the MN's primary path and returns an ASSCONF-ACK acknowledgement to the MN.
5. The MN's previous primary address is deleted when the ASCONF (Delete) message is sent to the CN.
6. The CN deletes this address andsends a confirmation message (ASSCONF-ACK) to the MN.

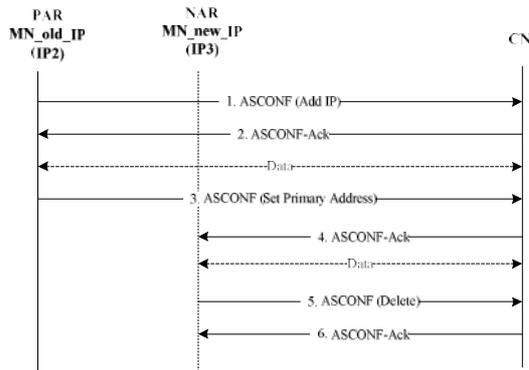The HTM^{global} handover procedure is illustrated in Fig. 9.

Figure 9: HTM^{global} handoff procedure

4   ANALYTICAL MODEL

To study the effectiveness of the proposed HTM, our comparison will consider the mSCTP handoff procedure illustrated in Fig. 1 since it is, to the best of our knowledge, the only procedure adopted in the previous mSCTP-based mobility proposals. The conducted analysis focuses on signaling cost, handoff latency and packet loss.

4.1 Preliminary and notations

Fig. 10 illustrates a typical mobility scenario where an MN starts its movement from the X_{start} point and ends at the X_{end} point. During its movement, an MN can perform either handoffs of type (a) or type (b) as indicated in Fig. 10. Handoff of type (a) refers to inter AMU domain handover (i.e., local handoff). A handoff of type (b) refers to the end-to-end handover performed outside an AMU domain (i.e., global handoff).

Let μ_r be the border crossing rate of an MN through access routers (ARs),

Let μ_d be the border crossing rate of an MN through AMU domains,

Let μ_I be the border crossing rate through ARs when the MN remains inside an AMU domain, μ_I is defined as: $\mu_I = \mu_r - \mu_d$.
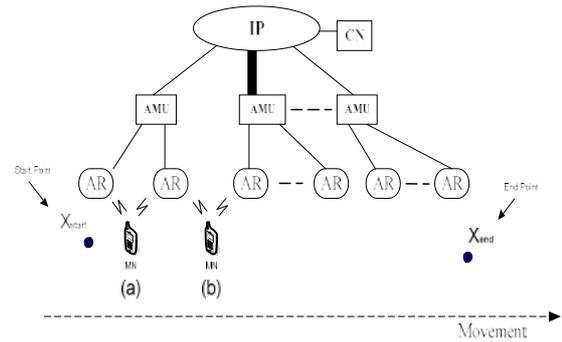

Figure 9:   MN roaming topology

According to [2], if we assume that an AMU coverage area is composed of M circular access router subnets, the border crossing rates can be expressed as:

$$\mu_d = \frac{\mu_r}{\sqrt{M}} \qquad (1)$$

$$\mu_I = \mu_r \cdot \frac{\sqrt{M}-1}{\sqrt{M}}$$

Based on the aforementioned work, μ_r can be defined as: $\frac{\rho \cdot v \cdot R_s}{\pi}$, where: ρ is the user density, v the MN average velocity and R_s the perimeter of a subnet.

In order to study the effectiveness of the proposed mobility mechanism we consider a traffic model composed of two levels, a session and packet. The MN mobility will be modeled by the cell residence time and a number of random values introduced in [3]. Generally, we model the incoming sessions as a Poisson process (i.e., inter-session arrival time are exponentially distributed). According to [3], the inter-session arrival time may not be exponentially distributed. Thus, alternative distribution models such as Hyper-Erlang, Gamma and Pareto have been proposed. However, performance analyses show that the exponential approximation remains an acceptable tradeoff between complexity and accuracy [3]. Therefore, for simplicity we assume that the MN residence time in an AR subnet and in an AMU domain

follow exponential distribution with parameters μ_r and μ_d respectively, while session arrival process follows a Poisson distribution with rate λ_s. Hence, if we denote: $E(N_r)$ as the average number of AR subnet crossing, $E(N_d)$ as the average number of AMU domain crossing and $E(N_I)$ as the average number of AR subnet crossing performed inside an AMU domain, we can define the above averages as introduced in [21] by:

$$E(N_r) = \frac{\mu_r}{\lambda_s} \quad (2)$$

$$E(N_d) = \frac{\mu_d}{\lambda_s} \quad (3)$$

$$E(N_I) = \frac{\mu_I}{\lambda_s} \quad (4)$$

Notations used in our analysis are given in Table 1.

Table 1: Notation

$T_{X,Y}$	transmission cost between node X and node Y
P_Z	processing cost at node Z
$N_{hop}^{X,Y}$	number of hops between node X and Y
δ	a proportionality constant to illustrate that the transmission cost for wireless hops are superior to those of wired hops
T_{hop}^c	transmission cost per hop
l_X	one lookup cost at node X
η_X	packet tunneling cost at node X
$D_{X,Y}$	transmission delay between nodes X and Y
$D_{tunneling}$	packet tunneling time
P_Z^t	processing time at node Z
T_{MD}	Movement Detection delay
T_{AC}	Address Configuration delay
T_{L2}	L2 handoff delay
T_{UF}	AMU Update and packet Forwarding delay

In what follows, we use the above equations to analyze both signaling and packet delivery costs of the studied mobility schemes.

4.2 Total cost analysis

We define the total cost (C_{total}) as the sum of signaling and packet delivery costs. In other words, C_{total} is given by:

$$C_{total} = C_{signal} + C_{delivery} \quad (5)$$

The signaling cost refers to the amount of signaling traffic while the packet delivery cost refers to the network overhead. The C_{signal} and $C_{delivery}$ are modeled during an inter-session arrival time that refers to the interval time between the arrival of the first packet of a data session and the arrival of the first packet of the next data session (i,e., one session lifetime). Note that signalling cost required for L2 handoff and address configuration are not

considered in our analysis since they are the same for the compared protocols.

4.2.1 HTM total cost

The HTM total cost is defined as:

$$C_{total}^{HTM} = C_{signal}^{HTM} + C_{delivery}^{HTM} \quad (6)$$

- HTM signaling cost

The HTM signaling cost is incurred when an MN performs either local or global handoffs. This cost is given by:

$$C_{signal}^{HTM} = E(N_I) \cdot C^{AR} + E(N_d) \cdot C^{AMU} \quad (7)$$

Where :

C^{AR} : refers to the signaling cost when an MN performs a handoff of type (a)

C^{AMU} : refers to the signaling cost when an MN performs a handoff of type (b)

Moreover, if we assume that a handoff preparation is always followed by a handoff execution, the expressions relevant to C^{AR} and C^{AMU} are given in Table 2.

Table 2: Expression of signalling costs

C^{AR}	=	$T_{MN_p,AMU} + T_{MN_n,AMU} + 2 \cdot T_{MN_n,CN} + 2 \cdot P_{AMU} + P_{CN}$
C^{AMU}	=	$3 \cdot T_{MN_p,CN} + 3 \cdot T_{MN_n,CN} + 3 \cdot P_{CN}$

MN_p and MN_n refer respectively to the MN's location before and after a handoff. The $T_{X,Y}$ cost can be expressed as:

$$T_{X,Y} = (N_{hop}^{X,Y} - 1 + \delta) \cdot T_{hop}^c \quad (8)$$

To illustrate the impact of the MN's mobility and the MN's average session arrival on the HTM signaling cost, we introduce a session-to-mobility factor (SMR) which represents the relative ratio of session arrival rate to the mobility rate.

The SMR factor is expressed by : $SMR = \frac{\lambda_s}{\mu_r}$ (9).

Hence, if we consider equations (1), (4) and (9), the equation (7) becomes:

$$C_{signal}^{HTM} = \frac{1}{SMR\sqrt{M}} \left[(\sqrt{M} - 1)C^{AR} + C^{AMU} \right] \quad (10)$$

- HTM packet delivery cost

Let A_p be the average packets sent by the CN during one session lifetime. Based on Fig. 11, the MN can perform either handoffs of type (a) or (b). However, only handoffs of type (a) incur a table lookup and an IP tunneling costs at the AMU. Hence the HTM packet delivery cost is given by :

$$C_{delivery}^{HTM} = A_p \cdot T_{MN,CN} + E(N_I) \cdot (l_{AMU} + \eta_{AMU}) \cdot A_p^{(a)} \quad (11)$$

Where: $A_p^{(a)}$ refers to the average packet tunneled during handoffs of type (a),

4.2.2 mSCTP total cost

The mSCTP total cost is defined as:

$$C_{total}^{mSCTP} = C_{signal}^{mSCTP} + C_{delivery}^{mSCTP} \quad (12)$$

- mSCTP signaling cost

Based on the mSCTP handoff procedure given in Fig. 9, the mSCTP signaling cost is given by:

$$C_{signal}^{mSCTP} = (E(N_l) + E(N_d)) \cdot (3 \cdot T_{MN_p,CN} + 3 \cdot T_{MN_n,CN} + 3 \cdot P_{CN}) \quad (13)$$

To express equation (13) as a function of the SMR factor, we use equations (1), (4) and (9).

$$C_{signal}^{mSCTP} = \frac{3}{SMR} \cdot (T_{MN_p,CN} + T_{MN_n,CN} + P_{CN}) \quad (14)$$

- mSCTP packet delivery cost

Since the mSCTP handoff procedure did not

incur any IP tunneling or table lookup costs, its packet delivery is given by:

$$C_{delivery}^{mSCTP} = A_p \cdot T_{MN,CN} \quad (15)$$

4.3 Handoff latency and packet loss

The handoff latency is defined as the time elapsed between sending of the last data packet through the old MN's primary address (i.e., old location) and receiving the first data packet on the MN's new primary address (i.e., new location). The packet loss refers to the amount of packets lost during this disruption time.
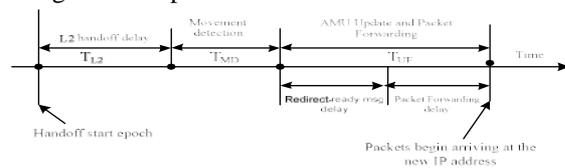
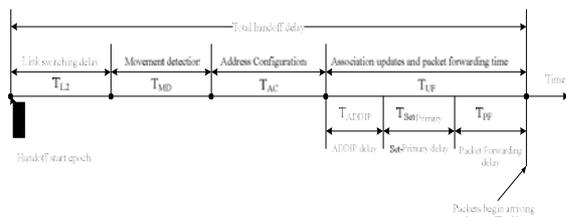Figure 10: HTM local handoff timeline delay

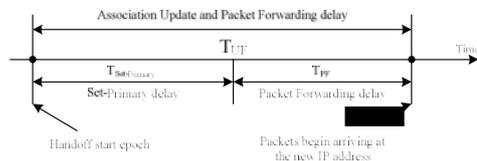Figure 11: mSCTP horizontal handoff timeline delay

Figure 12: mSCTP/HTM vertical handoff timeline delay

If a mobile node moves through cells belonging to a same technology (horizontal handoff), it cannot simultaneously use its two interfaces since it needs two transceivers according to the majority of radio systems [15]. However, if it performs a handover between heterogeneous wireless technologies (i,e., vertical handoff), it can use its interfaces in parallel. This means, that the MN continues to receive traffic on its old path

while it performs L2 link switching, movement detection, address configuration through the new interface and the association update (ADDIP). Practically, we can divide handoff latency into: link switching or L2 handoff delay (T_{L2}), movement detection delay (T_{MD}), address configuration delay (T_{AC}) and association updates and packet forwarding time (T_{UF}).

According to the handoff scenarios depicted in Fig. 11, an MN can perform either handoffs of type (a) or (b). Hence, we define the average handoff latency for HTM as:

$$D_{handoff}^{HTM} = \frac{1}{E(N_l) + E(N_d)} \left[(_l) \cdot EN \cdot D_{handoff}^{(a)} + EN \cdot P \cdot D (_h \cdot D_{handoff}^{(b),horizontal} + (1 - _h) \cdot D_{handoff}^{(b),vertical}] \quad (16)$$

Where:

$D_{handoff}^{(a)}$: latency relevant to handoff of type (a) (i.e., inside an AMU domain), the corresponding timeline delay is given in Fig. 12.

$D_{handoff}^{(b),horizontal}$: latency relevant to horizontal handoff performed outside an AMU, the corresponding timeline delay is given in Fig. 13.

$D_{handoff}^{(b),vertical}$: latency relevant to a vertical handoff, the corresponding timeline delay is given in Fig. 14.

P_h : probability that an MN perform a horizontal handoff outside an AMU domain.

The expressions of $D_{handoff}^{(a)}$, $D_{handoff}^{(b),horizontal}$ and $D_{handoff}^{(b),vertical}$ are given in Table 3.

Table 3: Expression of HTM handoff delays

$D_{handoff}^{(a)}$	=	$T_{L2} + T_{MD} + 2 \cdot D_{MN,AMU} + D_{tunneling} + P_{AMU}^t + \tau$
$D_{handoff}^{(b),horizont}$	=	$T_{L2} + T_{MD} + T_{AC} + 4 \cdot D_{MN,CN} + P_{CN}^t + \tau$
$D_{handoff}^{(b),vertica}$	=	$2 \cdot D_{MN,CN} + P_{CN}^t + \tau$

If we consider that $\mu_d \neq 0$ (i,e., we have at least two AMU domains), we use equations (3) and (4) to derive the following relation:

$$D_{handoff}^{HTM} = \frac{1}{\sqrt{M}} \cdot \left[(\sqrt{M} - 1) \cdot D_{handoff}^{(a)} + P_h \cdot D_{handoff}^{(b),horizontal} + (1 - P_h) \cdot D_{handoff}^{(b),vertical} \right] \quad (17)$$

Where: τ refers to the time between the instant when the sender is ready to send data packets and the instant when it effectively starts sending data packets to the MN's new location

According to [16] $D_{X,Y}$ is defined as:

$$D_{X,Y} = \frac{1-q}{1+q} \cdot \frac{s}{B_{wl}} + L_{wl}) + (N_{hop} - 1) \cdot (\frac{s}{B_w} + L_w + \varpi_q) \quad (18)$$

Where s is the message size, ϖ_q is the average queuing delay at each intermediate router, q is the probability of wireless link failure, B_{wl} (resp B_w) the bandwidth of wireless (resp wired) link and

L_{wl} (resp L_w) wireless (resp wired) link delay.

With mSCTP, the handoff latency is given by:

$$D_{handoff}^{mSCTP} = \frac{1}{\sqrt{M}}\left[\left(\sqrt{M} - 1 + P_h\right) \cdot D_{handoff}^{b,horizontal} + (1 - P_h) \cdot D_{handoff}^{b,vertical}\right] \quad (19)$$

On the other hand packet loss is proportional to the handoff delay since all data packets exchanged during this disruption period are lost. Practically, let λ_p be the packet arrival rate, the packet loss for both HTM and mSCTP is defined as:

$$\begin{cases} P_{loss}^{HTM} = \lambda_p \cdot D_{handoff}^{HTM} - Min(B_{HTM}, B_{AMU}) \\ P_{loss}^{mSCTP} = \lambda_p \cdot D_{handoff}^{mSCTP} \end{cases} \quad (20)$$

Where, B_{HTM} is the buffer size required for HTM and B_{AMU} the buffer size available at the AMU. The buffer size required for HTM is proportional to packet arrival rate and it is computed as follows:

$$B_{HTM} = \lambda_p \cdot (T_{L2} + T_{MD} + T_{UF}) \quad (21)$$

5 PERFORMANCE ANALYSIS

This section presents simulation and numerical results obtained when an MN uses either the proposed HTM or the mSCTP based handoff procedure. We choose mSCTP as the benchmark transport layer mobility protocol for our comparison since all the previous SCTP-based mobility proposals use the mSCTP standard. Moreover, mSCTP is a general IETF purpose standardized under the RFC 5061.

5.1 Simulation Setup

The main concern of our simulations is to show how the introduced AMU unit improves handoff seamlessness. That is why we consider the simulation scenario depicted in Fig. 14. This scenario is designed in such a way to provide realistic results, while remaining sufficiently small to be handled efficiently with the ns-2 simulator [34]. Simulation code is based on the SCTP module developed at the University of Delaware [35]. This SCTP module is modified so that it can support the newly introduced ADDIP-Soft Chunks, as well as AMU functionalities (Section III).

The MN is supposed to be single-homed since we will focus on local handoffs. Initially, the MN is assigned to AR1 and benefits from an ongoing association with CN. When the MN moves from AR1 to AR2, it performs a local handoff (inside an AMU). In all simulations, the observed MN moves at various speeds, on a straight line, from AR1 to AR2 sub-network. Each AR operates according to the 802.11b (11 Mbit/s) standards in the Distributed Coordination Function (DCF). Delays for both 802.11b WLANs equal 15 ms. A CBR agent is attached to CN and the MN operates as a sink. The average experiment time lasts around 300 s.
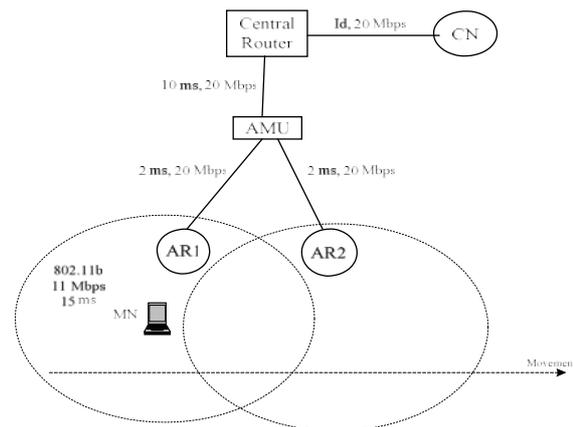

Figure 13: Simulation network topology

5.2 Simulation Results

Fig. 15 illustrates handoff latency behavior when an MN completes HTM^{local} and mSCTP handoffs. In fact, several experiments were conducted where the MN performs a handoff from AR1 to AR2, then it returns back to AR1. In each experiment, a wired hop is added between the MN and the CN, meaning that an additional delay is added to the CN-AMU link. The first thing to be noted is that when the number of intermediate hops between the MN and the CN increases, the mSCTP latency values continue to increase, while HTM^{local} latency remains approximately constant. This situation is due to the fact that HTM^{local} uses the AMU unit to redirect packets to the MN's new location as quick as possible. Then, it updates its association. This approach is completely different from mSCTP that has to update the MN's active association with ADDIP and Set-Primary chunks during the disruption time. Moreover, the HTM^{local} handoff latency remains lower than mSCTP one even if the distance between MN and CN is low. Indeed, with HTM^{local}, the MN anticipates its address configuration process by using the AMU unit. Obviously, this feature cannot be completed by mSCTP. Recall that the address configuration delay may take over than 500 ms [25].
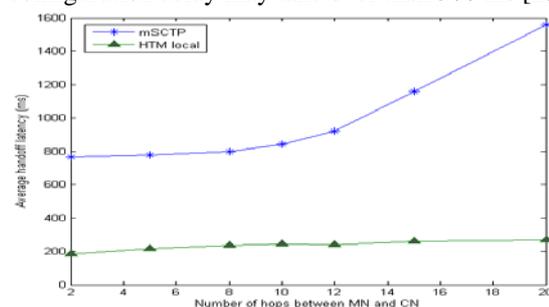

Figure 14: Impact of on MN-CN distance on handoff latency

Fig. 16 presents the average handoff latency, for both mSCTP and HTM, as a function of moving speed. Here, we set id = 20ms (i,e., delay between central router and CN) and we increase the MN's speed (v) from 2 m/s to 30 m/s while it performs several handoffs starting from AR1 to AR2. We notice that when the MN's speed is small, HTM shows lower handoff delay than mSCTP. However, when v > 12 m/s, the HTM's latency increases considerably and becomes equivalent to the mSCTP one. This is because when the moving speed increases, the sojourn time in the overlapping area becomes too small, so the MN do not have enough time to perform its configuration process. Therefore, the advantage of introducing the AMU unit is no longer considered when the MN's moving speed is high.
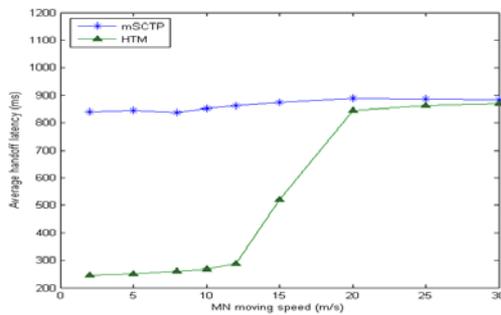

Figure 15: Impact of moving speed on HTM / mSCTP latencies

To illustrate how the proposed HTM improves throughput, consider the results illustrated in Fig. 17. These results correspond to the throughput relevant, respectively, to the previous and new MN's paths, i.e, the MN changes its point of attachment from AR1 to AR2. Observe that the throughput of the previous path decreases during the time interval $t \in [13s, 25s]$ where the handoff takes place. This drop is due to the increasing loss rate of AR1 as the MN moves. Once the handoff is over, notice that the MN throughput increases again until it reaches its original level. However, the throughput reported immediately after the handoff remains lower than the one computed before the handoff occurred.
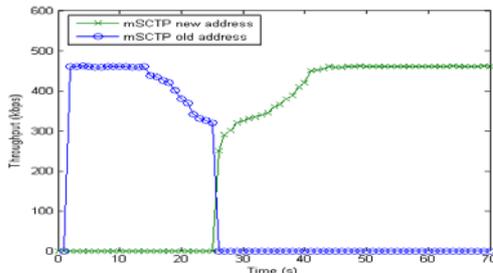

Figure 16: Throughput relevant to an mSCTP path handover

This situation is due to failed SACKs that cause a

diminution of the congestion window (CWND), thus reducing throughput. To show how the proposed HTM improves throughput compared to mSCTP, consider the throughput obtained immediately after a handoff for HTM and mSCTP.

Fig. 18 shows the throughput pertaining to the time interval (25-40s) following an MN handoff. Note that the HTM throughput is relatively high compared to mSCTP. This is due to the fact that HTM [local] uses the AMU unit to buffer and forward all the traffic to the new MN's location. This traffic obviously includes SACKs which are not lost, unlike what happens with mSCTP. Accordingly, MN will receive a majority of its SACKs within the RTO time interval (Retransmission TimeOut) since the HTM_{local} latency is less than 300 ms while the RTO interval is about 1 second.
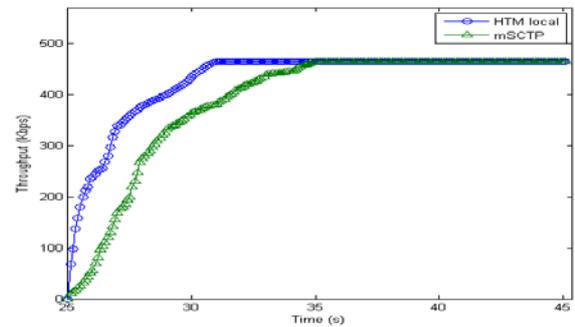

Fig. 17. Throughput of HTM [local] vs mSCTP

5.3 Numerical Results

In this section, we use the developed cost models (section 4) to illustrate how the proposed mobility scheme HTM improves QoS parameters in terms of signaling cost, handoff delay and packet loss compared to mSCTP.

The list of the parameter values used for our numerical results is shown in Table 4.

Table 4: Parameters used for performance analysis

Parameters	Symbols	Values
Wireless link failure probability	q	0.5
Average queuing delay	ϖ_q	0.1 ms
Wired link delay	B_w	100 Mbps
Wireless link bandwidth	B_{wl}	11 Mbps
Message size	s	296 bytes
Number of AR subnets per AMU/MAP domain	M	4
Average packet arrival per session	A_p	20
Average packets tunneled during a handoff of type (a)	$A_p^{(a)}$	2
Lookup cost at the AMU	l_{AMU}	2
Packet tunneling cost at the AMU	η_{AMU}	2
L2 handoff delay	T_{L2}	50 ms
Movement detection delay	T_{MD}	100 ms
Address Configuration delay	T_{AC}	500 ms
Waiting time before effective data transmission	τ	1 ms

Fig. 19 illustrates the total signaling cost as a function of the SMR ratio. When the SMR ratio is inferior to 1, the mobility rate is higher than the session arrival rate that is why the signaling cost increases for both HTM and mSCTP. This increase becomes more noticeable when the SMR is close to 0. However, the HTM cost remains lower than the mSCTP cost. On the other hand when the SMR is superior to 1, i,e., the session arrival rate is greater than the mobility rate, the binding updates, relevant to handoffs, are performed less often.
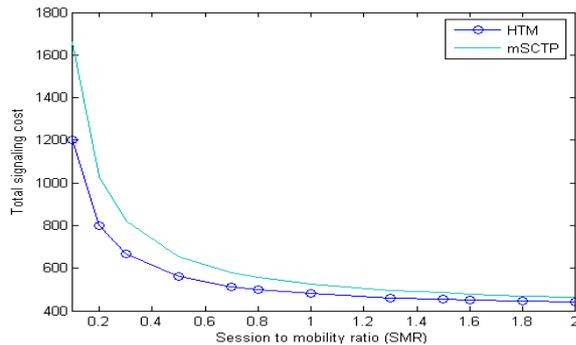

Figure 18:   Impact of the SMR on the total signaling cost

Fig. 20 illustrates the total signaling cost as a function of mobile node velocity. We notice that the total signaling cost increases for both HTM and mSCTP. However, the signaling costs involved by HTM remain lower than mSCTP. Moreover the gap between mSCTP and HTM signaling costs becomes more important when the MN's velocity increases. This behaviour is to be expected since the MN will perform frequent handoffs when its velocity reaches high values. Nevertheless, HTM reduces the amount of signalling cost since it takes into account local handoffs.
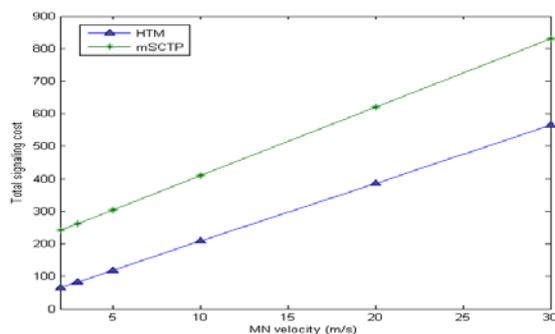

Figure 19:   Impact of the MN velocity on the total signaling cost

Fig. 21 shows that the HTM total signaling cost is proportional to the AMU tunneling cost. However, it remains lower than the mSCTP cost even if high values are used for the AMU tunneling cost (i., more that 20). Recall that all of the processing costs used for our performance analysis are less or equal to 2. On the other hand, mSCTP is

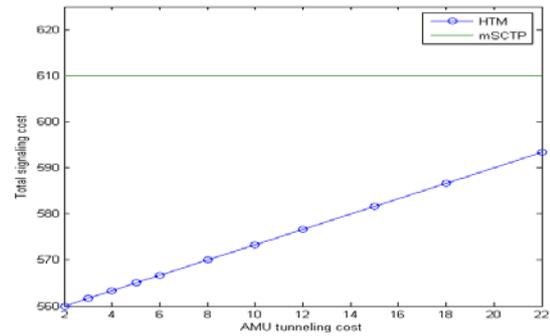not affected by the AMU cost variation since it does not perform traffic redirection.

Figure 20:   Impact of the AMU tunneling cost on the total signaling cost

In Fig. 22 we present the average handoff delay as a function of the wireless link delay. We notice that the average handoff delay is proportional to the wireless link delay for both HTM and mSCTP. However, it can be noticed that the HTM average latency is lower than mSCTP. Moreover, when P_h increases (i,e., probability of horizontal handoff performed outside an AMU unit), handoff latencies increase for both HTM and mSCTP. However, the HTM's latency remains lower than the mSCTP one. This means that the introduction of the AMU unit improves considerably the MN's handoff delays during its roaming through homogeneous networks.
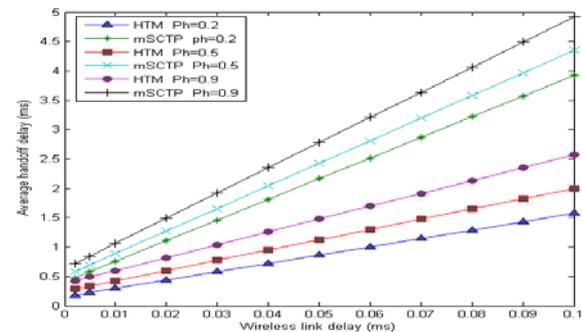

Figure 21:   Handoff latency as a function of wireless link delay

The impact of HTM [local] on the MN's latency is clearly illustrated in Fig. 23 where we compare the two scenarios of mSCTP handoffs (i.e., horizontal and vertical) with our proposed mobility scheme. Recall, that HTM (i.e., HTM[global]) and mSCTP use the same vertical handoff procedure. We notice that HTM [local] presents lower average handoff latency compared to mSCTP. The gap between HTM [local] 's latency and mSCTP becomes more and more important as the wireless link delay increases. This difference is particularly due to the absence of address configuration delay in HTM [local] . Moreover, the consideration of local mobility reduces considerably the association update delay since the CN is notified only when the MN is attached to its new location.
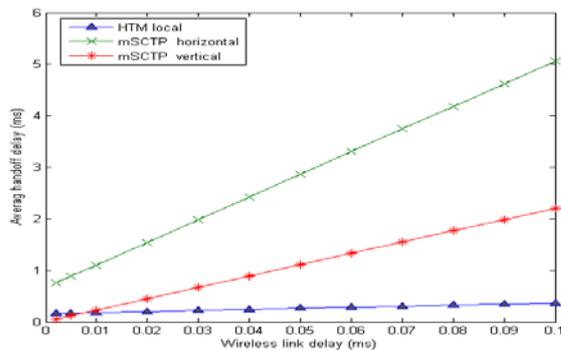
Figure 22: Impact of wireless link delay on horizontal/vertical handoffs

Figure 24: Packet loss behavior for different packet arrival rates

Fig. 24 illustrates handoff latency as a function of the average subnet crossing rate inside an AMU ($E(N_I)$). When this rate is low, i,e., most of the MN's handovers are performed in the absence of the AMU units, we notice that the average HTM latency is high. With the increase of $E(N_I)$, we observe a noticeable decrease of the HTM average latency which becomes approximately constant when this rate reach high values. This situation shows again that the consideration of local handoffs by our mobility proposal reduces considerably the overall average handoff delay when the MN performs consecutive horizontal and vertical handoffs. On the other hand, the mSCTP handoff latency remains high and almost insensitive to the $E(N_I)$ rate.
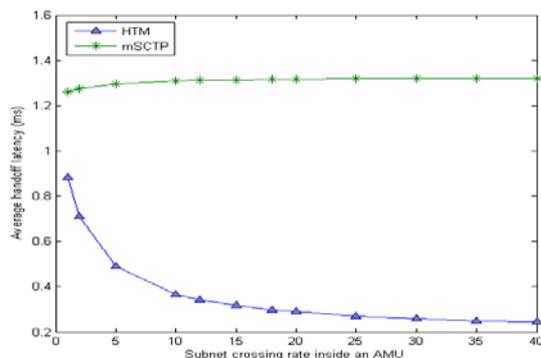

Figure 23: Impact of the intra-subnet crossing rate on the handoff delay

Fig. 25 shows the behavior of packet loss as a function of packet arrival rate. It is noticed that packet loss increases for both HTM and mSCTP. However, the HTM packet loss remains lower than mSCTP. This situation is quite normal since the handoff delays for HTM is lower than mSCTP and by definition of all of the packets received at this period are lost. In addition, HTM uses a buffering strategy when an MN roams inside a same AMU domain which helps to avoid as much as possible packet loss during the MN's disruption time.
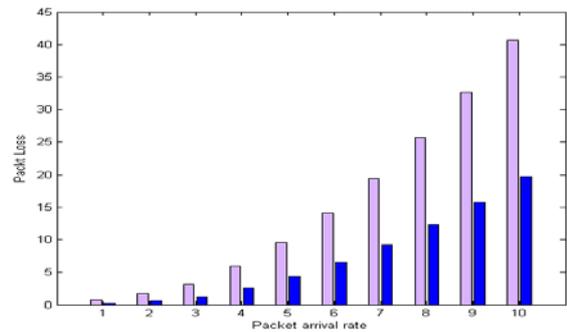
6   CONCLUSIONS

This paper proposes a new hierarchical transport layer mobility scheme called HTM, whose main goal is to provide mobile nodes with seamless roaming through heterogeneous networks. More specifically, HTM consists of an end-to-end mobility protocol based on SCTP features, which includes multihoming and ADDIP Extension. It particularly introduces an Anchor Mobility Unit (AMU) to deal with local mobility in order to reduce handoff latency and signaling load. Additionally, HTM addresses the problem of spurious retransmissions due to failed SACKs. Finally, to ensure mobile node tracking when initiating associations, a location management scheme that uses the Dynamic DNS service (DDNS) is introduced. Simulations and numerical results show that HTM ensures low latency, good throughput and limited signaling load compared to the mSCTP based handoffs. Future work shall investigate how this proposal can be adapted to mobile ad hoc networks as well as the impact of the proposed location management scheme on system performance.

7   REFERENCES

[1]   Frattasi, S., Fathi, H., Fitzek, F. H. P., Prasad, R., Katz, M. D. 2006. "Defining 4G technology from the users perspective". IEEE Network, Vol. 20, No. 1, pp. 35-41.

[2]   Bauman, F. V., Niemegeers, I. G. 1994. "An Evaluation of Location Management Procedures". Proceeding Of 3rd Annual Int'l Conference Universal Personal Communications (UP' 94), September 1994, pp. 359-364.

[3]   Fang, Y. 2003. "Movement-Based Mobility Management and Trade Off Analysis for Wireless Mobile Networks". IEEE Transactions on Computers, Vol. 52, No. 06, pp. 791-803.

[4]   Akyildiz, I. F., McNair, J., Ho, J. S. M., Uzunalioglu, H., Wang, W. 1999. "Mobility Management in Next-Generation Wireless Systems". Proceeding of the IEEE, Vol. 87, No. 8, pp. 1347-1384.

[5]   Gundavelli, S., Leung, K., Devarapalli, V.,

Chowdhury, K., Patil, B., "Proxy Mobile IP", RFC 5213, August 2008.

[6] Johnson, D., Perkins, C., Arkko, J. 2004. "Mobility Support in IPv6", IETF RFC 3775.

[7] Soliman, H., Castelluccia, C., El-Malki, K., Bellier, L. 2005. "Hierarchical mobile IPv6 mobility management (HMIPv6)", IETF RFC 4140.

[8] Koodli, G. 2005. "Fast handovers for mobile IPv6". IETF RFC 4068.

[9] Jung, H. Y., Kim, E. A., Yi, J. W., Lee, H. H. 2005. "A scheme for supporting fast handover in hierarchical mobile IPv6 networks". ETRI Journal, Vol. 27, No. 6, pp. 798–801.

[10] Snoeren, A., Balakrishnan, H. 2000. "An end-to-end approach to host mobility". ACM MobiCom Boston, pp. 155-166.

[11] Handley, M., Schulzrinne, H., Schooler, E., Rosenberg, J. 1999. "SIP: Session Initiation Protocol", IETF RFC 2543.

[12] Wei, S. H., Huang, M. H., Long, W. 2005. "Research on the scheme supporting mobility of TCP application", Journal of Chongqing University of Posts and Telecommunication, Vol. 17, No. 1, pp. 117-120.

[13] Stewart, R. 2007. "Stream Control Transmission Protocol". IETF RFC 4960.

[14] Stewart, R., Xie, Q., Tuexen, M., Maruyama, S., Kozuka, M. 2007. "Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration". IETF RFC 5061.

[15] Atallah, J. G. and Ismail, M., "Future 4G front-ends enabling smooth vertical handovers", IEEE circuits and Device Magazine, Vol. 22, No. 1, 2006, pp. 6-15.

[16] McNair, J., Akyildiz, I. F., Bender, M. D. 2001. "Handoffs for Real-time Traffic in Mobile IP version 6 Networks". Proceeding of IEEE GLOBECOM, Vol. 6, pp. 3463-3467.

[17] Peˊrez-Costa, X., Torrent-Moreno, M., Hartenstein, H. 2003. "A performance comparison of mobile IPv6, hierarchical mobile IPv6, fast handovers for Mobile IPv6 and their combination". ACM Mobile Comp. and Comm. Rev, Vol. 7, No 4.

[18] Ramjee, R., Varadhan, K., Salgarelli, L., Thuel, S. R., Wang, S. Y., La Porta, T., "HAWAII: A domain-based approach for supporting mobility in widearea wireless networks", IEEE/ACM Transactions on Networking, Vol. 10, No. 3, June 2002, pp. 396-410.

[19] Campbell, A.T., Gomez, J., Valko, A.G., "An overview of cellular IP ", Wireless Communications and Networking Conference, WCNC 1999 IEEE, Vol. 2, pp. 606-610.

[20] Zhang, J., Chan, H., Leung V. 2007. "A SIP-based Seamless-Handoff (S-SIP) Scheme for Heterogeneous Mobile Networks". in Proceeding of the IEEE Wireless Communications and Networking Conference, Hong Kong, pp. 3946-3950.

[21] Xiao, Y., Pan, Y., Li, J. 2004. "Design and Analysis of location Management for 3G Cellular Networks". IEEE Transactions on parallel and distributed systems, Vol. 15, No. 04, pp. 339-349.

[22] Hsieh, H., Kim, K., Zhu, Y., Sivakumar, R. 2003. "A receiver-centric transport protocol for mobile hosts with heterogeneous wireless interfaces". ACM MobiCom, San Diego, pp. 1-15.

[23] Maltz, D., Bhagwat, P. 1998. "MSOCKS: An architecture for transport layer mobility". INFOCOM San Francisco, pp. 1037-1045.

[24] Huitema, C., " Multi-homed TCP ", draft-huitema-multi-homed-0.txt, 1995.

[25] Mishra, A., Shin, M., Arbaugh, W., "An empirical analysis of the IEEE 802.11 MAC layer handoff process", ACM SIGCOMM Computer Communication Review, 2003, Vol. 33, No. 2, pp. 93-102.

[26] Ma, L., Yu, F., Leung, V. C. M., Randhawa, T. 2004. "A new method to support UMTS/WLAN vertical handover using SCTP ". IEEE Wireless Communications, Vol. 11, No. 4, pp. 44-51.

[27] Fu, F. S., Atiquzzaman, M., Ma, L., Ivancic, W., Lee, Y., Jones, J. S., Lu, S. 2004." TraSH : A Transport Layer Seamless Handover for Mobile Networks". Technical Report: OU-TNRI-04-100.

[28] Fu, F. S., Ma, L., Atiquzzaman, M., Lee, Y., "Architecture and performance of SIGMA: A seamless mobility architecture for data networks", 40th IEEE International Conference on Communication (ICC), Seoul, Korea, May 2005, pp. 3249-3253.

[29] Koh, S. J., Chang, M. J., Lee, M. 2004. "mSCTP for soft handover in transport layer". IEEE Communication Letters, Vol. 8, No. 3, pp. 189-191.

[30] Iyengar, J. R., Amer, P. D., Stewart, R. 2006."Concurrent Multipath Transfer Using SCTP Multihoming Over Independent End-to-End Paths". IEEE Transactions on Mobile Computing, Vol. 14, No. 5, pp. 951-964.

[31] Fracchia, R., Casetti, C., Chiasserini, C. F., Meo, M. 2007."WiSE: Best-Path Selection in Wireless Multihoming Environments". IEEE Transactions on Mobile Computing, Vol. 6, No. 10, pp. 1130 – 1141.

[32] Droms, R. 1997. "Dynamic Host Configuration Protocol", IETF RFC 2131, March 1997.

[33] Thomson, S., Narten, T. 1998. "IPv6 Stateless Address Autoconfiguraton". IETF RFC 2462.

[34] http://www.isi.edu/nsnam/ns/

[35] SCTP Module http://pel.cis.udel.edu/