Performance Evaluation of Mobile IP Protocols in a Wireless Environment

Maurizio Dell'Abate
Martino De Marco
Vittorio Trecordi
CEFRIEL / Politecnico di Milano, Milan, Italy

Abstract

Great interest in recent years has been devoted to mobile communications, as well as to Internet and IP based networks. The most promising protocol proposals for the handling of host mobility in IP networks are Mobile IP (MIP) and Route Optimization Mobile IP (ROMIP). The two schemes have been contrasted in the case of wireless mobile hosts linked by means of a wired backbone. The simulation results show that the performance in terms of packet delay and loss as well as the overhead of control traffic largely depend on the offered traffic pattern.

1 Introduction

Host mobility is becoming an increasingly important feature with the development of notebook and portable computers. Recent advances in wireless network interfaces and the implementation of the global networks make host mobility an issue of interest both in wireless and wired networks. This paper describes and compares two proposals for the support of mobility in IP networks, namely the Mobile IP basic protocol (MIP, ref. [1]) and the Route Optimization Mobile IP (ROMIP, ref. [2]). In particular, we address the performance of both proposals by simulation in the particular case of wireless mobile hosts linked by means of a wired backbone. The outline of the paper is as follows. In Section 2 the mobility in Internet scenario is introduced and MIP and ROMIP are presented. Section 3 presents the motivation of the work and the objectives of the paper. Section 4 presents the assumptions and the model adopted for simulating the system and drawing performance results. The numerical results are discussed in Section 5, while conclusions are left to Section 6.

2 Solutions for Internet Mobility

2.1 Mobility Requirements and Specifications

We define *host mobility* as the movement of the *local network interface* of a host over different Internet subnets. These subnets provide the datalink interface to mobile host, and are based either on wired technology (such as Ethernet or ATM) or on wireless media (such as radio-frequency LANs, infrared LANs or cellular networks).

The basic *Internet Protocol* (IP) network paradigm does not support host mobility. IP addresses are assigned to network interfaces in dependence on their physical location: in fact, the first field of an IP address (namely the NETID) is common to all the interfaces that are linked to the same subnet. This scheme enables route aggregation and thus hierarchical routing, but prevents any host from keeping its address while moving over different subnets.

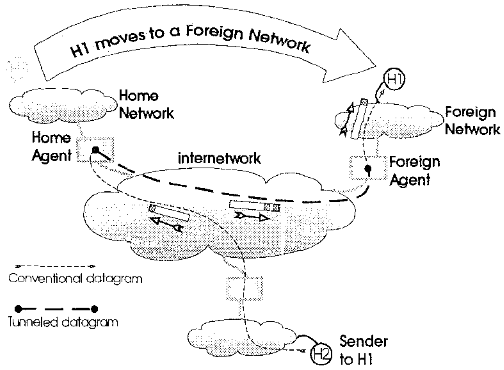

Figure 1: Basic triangle routing scheme in Mobile IP (MIP).

Therefore, IP should be enhanced in order to support host mobility. Applications need to use IP addresses as *location independent* identifiers. Any sender should transparently send IP datagrams to a given destination host independently on the location of the destination host. In order to achieve the delivery of packets to the addressed mobile hosts, *location tracking* and *handover* mechanisms must transparently operate at the network layer of the OSI protocol reference model. The latter features should guarantee support to *on-line mobility*.

The Mobile IP working group of the Internet Engineering Task Force (IETF) is specifying an enhanced version of IPv4, called Mobile IP, designed to offer transparent support to host mobility. After a long discussion phase between IETF and other research labs, a set of RFC documents has been finally published, thus placing Mobile IP proposal on the Internet standards track (ref. [1], [3], [4], [5]).

2.2 Mobile IP basic Protocol (MIP)

In this paragraph we present an enhancement to IPv4 enabling host mobility with light complexity (ref. [1]).

The MIP proposal refers to some architectural entities, namely:

- **Mobile Host:** A host that changes its point of attachment among different subnets. Notwithstanding its location, a host is always identified by its *home address* (that is, an IP address assigned in the home network's address space).
- **Home Agent:** A router on a mobile host's home network which stores the current location of the mobile host (*mobility binding*), in order to properly redirect incoming packets addressed to that host.

0-7803-4788-9/98/$10.00 © 1998 IEEE.

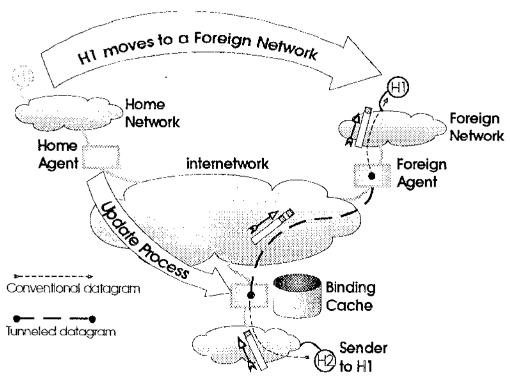

Figure 2: *Route Optimization scheme in Mobile IP (ROMIP).*

- **Foreign Agent:** A router on a mobile host's visited network which receives redirected packets and delivers them to the destination.

Mobility agents advertise themselves on the local subnet by periodically sending *Agent Advertisement* messages; these beaconing packets enable a mobile host to learn the address of the subnet it is linked to.

If a mobile host is on its home network, it sends and receives packets according to conventional IP mechanisms. As a mobile host comes to know that it is hosted on a foreign network, it obtains a *care-of address* in the local address space. The latter address may be the foreign agent's address or a dynamic address obtained by some external mechanism such as the Dynamic Host Configuration Protocol (DHCP) (ref. [6]). After having obtained a care-of address, a mobile host sends a *Registration Request* to its home agent, in order to inform it about its new location. A *Registration Reply* message is sent back by the home agent to acknowledge the binding update.

Packets sent to the mobile node's home address are routed to the home network by conventional IP mechanisms (see Fig. 1 for a routing example). Once to the home network, packets are intercepted by the home agent and then *tunneled* (ref. [3], [4], [7]) to the location of the mobile host by using the care-of address.

2.3 Route Optimization (ROMIP)

The basic MIP proposal has been enhanced by the ROMIP scheme (ref. [2]).

According to ROMIP hosts cache address bindings, whenever an address binding is cached, the packets addressed to that particular host are tunneled directly bypassing the destination home agent intervention (see Fig. 2 for a routing example). Some mechanisms are also provided in order to enable packets tunneled according to an out-of-date binding to be recovered and forwarded to the updated location.

ROMIP introduces custom control packets[1] in order to manage the binding caches in either fixed or mobile nodes, and to keep them lined up with current locations.

The complete description of ROMIP operation is quite complex, owing to a broad class of events that may trigger the sending of a control message. However, a more complete presentation of ROMIP messages and events will be given in the section dedicated to the simulation model, with reference to the protocol subset we addressed.

3 Motivation and Goals

Both Mobile IP protocol proposals should guarantee full operational transparency in the support to host mobility, thus enabling IP datagram service to be used exactly as if the destination nodes were linked to their home networks.

We are interested here in comparing the efficiency of the Mobile IP protocols (MIP and ROMIP), and in evaluating their performance transparency when the end-points of the traffic sessions[2] are applications running on mobile hosts. Therefore, some performance issues are now introduced and discussed, and the work goals are stated.

The basic Mobile IP protocol (MIP) apparently presents some inefficiencies:

- **Triangle routing:** The destination home agent is a fixed redirection point for exchanging every IP packet, even if a shorter routing path is available between source and destination. This can lead to unnecessarily large end-to-end packet delay.
- **Home Agent overloading:** Owing to triangle routing, the network links connecting a home agent to the network are easily overloaded. Indeed, all session paths sharing the subnet field of their destination address converge into that subnet home agent, even if adjacent network links are idle.

The above drawbacks are counterbalanced by some apparent advantages: the protocol is simple, exchange of control messages is limited, and the address bindings are highly consistent since they are kept in one single point for a given host (its home agent).

The enhanced Mobile IP protocol (ROMIP) apparently presents some efficiency advantages, owing to the availability of mechanisms that enable route optimization and caching of bindings. In the following items, some of these advantages are explained.

- **Direct routing:** ROMIP allows every traffic source to cache and use binding copies. The original binding for a mobile host is kept in its home agent, but ROMIP supports a further *update process* by which a binding copy can be propagated to the requiring nodes, which may keep it in their cache for immediate or future use. Local bindings enable most packets in a traffic session to be delivered by direct routing, with apparent gain in terms of quality of service and scalability. The above update process will be completely outlined in the model section.
- **Handover management:** A moving host always informs its previous foreign agent about the new care-of address, so that packets tunneled to the old location (owing to an out-of-date binding copy) can be forwarded to the current location. With MIP, those packets had to be discarded or sent to the home agent again. The above handover support should increase overall quality of service in case of high mobility.

The mentioned advantages are counterbalanced by some apparent disadvantages: the protocol is quite complex, exchange of control messages and processing overhead (due to cache queries) could be critical, and cached bindings are possibly inconsistent since they are kept in a distributed fashion.

Technical literature is quite poor of performance evaluation works oriented to routing efficiency in a mobile Internet context. A few studies have evaluated Mobile IP proposals previous to related IETF activities, but no previous work makes a

---

[1] We define as *user packet* a packet exchanged between two IP service's users, and as *control packet* a packet exchanged only between two IP layer entities.

[2] We define a *traffic session* as a protracted exchange of user packets between a source and a destination host.

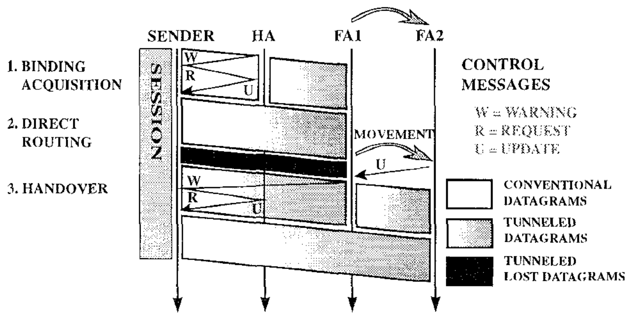

Figure 3: *A time model for Route Optimization protocol.*

comparison between MIP and ROMIP (ref. [8], [9]); moreover, explicit wireless assumptions have not been made in an internetwork model. There is also a number of numerical analysis of new backoff alghoritms wireless-optimized for use in TCP, but these topics don't concern network-layer functionality (for example, ref. [10]).

In this context, our work has two main goals: (i) to analyze quantitatively mobility effects on a Mobile IP network, and (ii) to evaluate the performance of the above outlined Mobile IP routing strategies, MIP and ROMIP.

Particularly, the evaluation of ROMIP enhancement requires to estimate quality of service parameters (end-to-end packet delay, packet loss due to mobility effects) and network-layer load parameters (transmission and processing overhead).

Although IP operation is actually independent from the underlying layers, we assume mobile hosts always obtain a dedicated bandwidth wireless connection to the currently visited subnet. All wireless access points are supposed to be connected by a fixed wired link pattern, giving rise to the simulated backbone internetwork. In the following section, the simulation model is more precisely outlined.

4 Hypothesis for Simulation

This section presents the model adopted as test-bed for the performance evaluation. The description will be given by presenting first the subset of protocol events included in our simulator. A time-based diagram describing our scheme of MIP and ROMIP operation is introduced and discussed. Some implementation choices are also stated, in those areas where the protocol does not specify a recommended behaviour.

We then provide a description of architectural aspects, such as physical and topological properties of the fixed network used as transport backbone. Finally, attention will be turned to the users of the Mobile IP service, that is to the parameters chosen to model (i) the service requested by the mobile users, (ii) their mobility pattern, (iii) the local area wireless transmission.

4.1 Update Process Model

The simulated ROMIP operation fully complies with the description previously given and to the IETF specification fundamentals. Now a temporal diagram for the protocol is shown, that depicts an example of working in the simulated context; the MIP operation model can be easily obtained as a subset of the ROMIP one. The control messages provided by ROMIP events will be introduced as they are involved.

Refer to Fig. 3 in the following outline. At time origin a tagged host starts sending user packets to a randomly selected mobile host, that is currently visiting the subnet whose foreign agent is named FA1. The vertical band placed on the left side of the diagram stands for a compact sequence of outgoing user packets, all having the same destination address. This packet stream models a constant-emission traffic session between two mobile hosts (as obvious, source mobility does not matter).

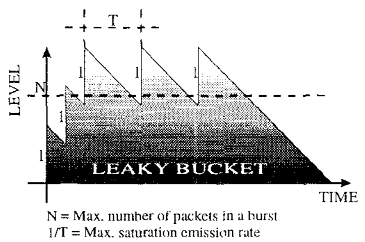

Figure 4: *Leaky bucket policy for control of warning emission.*

Three working phases are outlined, numbered from upper to lower side of the diagram. We start our description from the first phase, and follow increasing values of time towards the lower side of the diagram.

1. **Binding acquisition:** The packets sent in this phase reach the destination home agent (say HA) and they are then tunneled to the current foreign agent (say FA1). HA is supposed to have been previously informed of the destination's current location by the sending of a *binding update message* (not shown in the figure) from the mobile host.

   HA, just after having tunneled the first packet, sends a *binding warning message* (W) back to the source, informing the latter that it appears to have no valid binding. The source, in response to this warning, sends a *binding request message* (R) to the HA, keeping on sending user packets in the meanwhile. The HA replies with a binding update message (U), containing the requested care-of address.
2. **Direct routing:** The source caches the received binding and uses it to tunnel its packets directly to FA1; this phase continues steadily as far as the destination moves to another subnet.
3. **Handover:** The destination suddenly moves under another foreign agent (FA2); just after its movement, it sends two binding update messages (U), both to its home agent (HA) and to its previous foreign agent (FA1) (the former message is not shown in the diagram).

   The source has no way to get aware of the movement and keeps on emitting user packets to FA1. These packets get lost (see the black band in the diagram) until FA1 receives the above update.[3] As soon as FA1 gets updated, it warns the source (W) and forwards incoming packets to the actual location (FA2). The handover phase ends when the source, having received a fresh binding from HA, can enter a direct routing phase again.

The described cycle persists by repetition of phases 2 and 3, until the source closes the traffic session. This source, after an idle period, will select a new destination and a new cycle will start from the first phase.

Two implementation choices concerning the above model are now stated.

- A node that issues binding warning messages should control the rate at which they are sent to another node regarding a given mobile host. After a few warning emissions to

---

[3] In order to model time-out based expiring, we assume that a foreign agent doesn't remove a visiting host's entry just after the host's leaving, but that it does it later, when the host's update message is received. In the meanwhile, in-flight packets gets lost since they're transmitted on the wrong subnet.

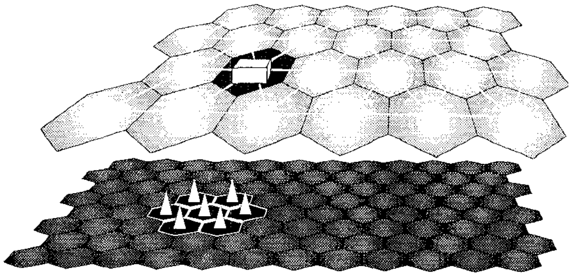

Figure 5: *Internet routing layer and datalink layer.*

the same node about the same mobile host, a back-off algorithm is used to limit the frequency of new notifications. It has been adopted a leaky-bucket policy to control warning emission (see Fig. 4).

- Control messages addressed to mobile hosts (such as binding updates) are always delivered by triangle routing, even if a valid binding for the destination is found in a transit node's cache. Leaving route optimization reserved to user packets avoids for further control messages to be issued in delivering a control message, possibly giving rise to positive feedback flooding phenomena.

4.2 Fixed Network Topology

As stated by the protocol authors (ref. [5]), Mobile IP was born essentially to give support to the *macro-mobility* problem, leaving the datalink entities to perform *micro-mobility* functions, such as link-layer handovers between wireless transceivers. Owing to the indefinite and time-variable structure of a wide-area internetwork, we have chosen a regular and homogeneous network to be our fixed backbone (ref. [11]). The corresponding graph is shown in the upper plane of Fig. 5. We deal with a flat lattice, having a parallelogram shape, whose linear dimensions are assigned (X and Y, node number for each side).

The graph nodes correspond to the network switching centers, that is the Mobile IP routers, performing datagram forwarding among the outgoing fixed links, according to the well-known connectionless hop-by-hop paradigm. The graph edges represent just these fixed links, that is the lines bearing datagrams between router couples; the fixed links are supposed to support bi-directional data traffic (a non-oriented graph is used).

We have chosen a flat, open territorial domain (not closed upon itself) for two main reasons: (i) our testbed can't refer to the global Internet, but only to an arbitrarily limited subset; (ii) if a closed torus network were adopted, the shortest routing path between two mobile users could change discontinuously after one end-point's movement.

Our topological model doesn't make provision for a hierarchical structure, owing to the lack of any functional aspect in the Mobile IP protocols dealing with mobility information aggregation. We may consider the graph nodes (i.e. the Mobile IP routers) as the gateways for a complex domain whose internal structure is left unspecified, except for the local presence of a number of mobile users, sources and drains of user traffic.

The upper plane in Fig. 5 is therefore the object of our analysis: in that structure the Mobile IP protocols are left working, providing datagram service to the mobile hosts applications. Each node in that plane characterizes an autonomous system, or a logical cell: these are atomic domains with respect to network-layer protocols and to the mobility process. The lower plane represents only an hypothetical datalink layer, neither modeled nor simulated, in which a cluster of base-stations provides the radio interface to the mobile stations (ref. [12] for an analogous model).

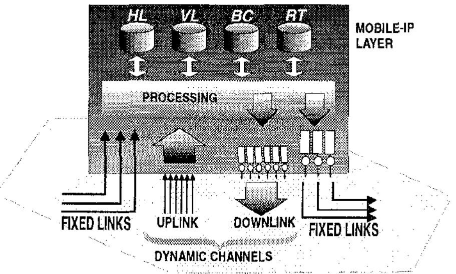

Figure 6: *Mobile IP router model.*

4.3 Mobile IP Router Model

Because the main function of a router is to forward autonomous unit of data (datagrams), the single-packet level of aggregation has been chosen to describe and implement information flows involved in the protocol models.

Each router is modeled as shown in Fig. 6. The processing module is fed by two sets of packet streams: the fixed link streams set, carrying packet traffic from linked nodes, and the radio uplink streams set, carrying packet traffic from mobile hosts currently registered in the local area. Each input stream should be considered as the output line of the waiting queue associated with the feeding entity.

The processing module scans each incoming packet, doing the appropriate action according to the simulated protocol. Incoming packets may be user packets or control packets, encapsulated or not.

Each router performs (i) the basic routing functionality (its routing table, RT, is precomputed according to Dijkstra's shortest path alghoritm); it performs (ii) home agent functionality for a predefined set of mobile hosts (its home list, HL, contains care-of addresses notified by binding update messages); it performs (iii) foreign agent functionality for the visiting mobile hosts (listed in the visitor list, VL); finally, for ROMIP only, it performs (iv) cache agent functionality (the binding cache, BC, stores recent binding updates and it is managed according to a *least recently used policy*, LRU).

4.4 Traffic and Mobility Patterns

In the simulated context, mobile hosts are the sources and the destinations for all traffic sessions (control messages, instead, may even originate or end in the fixed nodes).

In Fig. 7 a module decomposition for a mobile host is shown. The upper module implements application-layer functions, accessing IP datagram service to send or receive user packets (these flows are shown by white arrows). The underlying Mobile IP layer sends (receives) user packets to the uplink (from the downlink) dedicated channel. It encapsulates leaving packets if a valid binding is found in the binding cache (BC), and it decapsulates the received packets delivered by tunnel, passing them to the upper layer if they are of user type. Moreover, the control messages provided by ROMIP are sent, received or processed by Mobile IP processing module: these control flows (shown by black arrows) are not visible to application, exactly as in the real implementations.

The uplink and the downlink lines model a bi-directional dedicated wireless channel with preassigned bandwidth (as in cellular networks). This wireless channel links a mobile host to

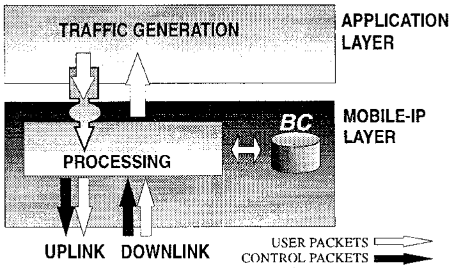

Figure 7: *Mobile host model.*

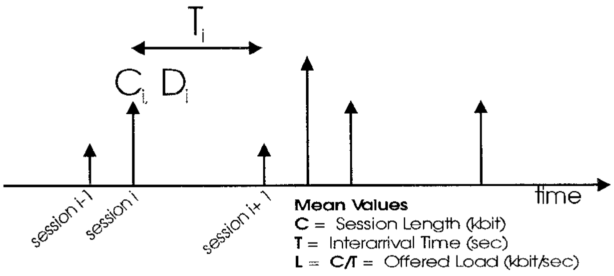

Figure 8: *User traffic generation process.*

the local router, which is supposed to be able to simultaneously access, *without blocking*, all the downlinks to the visiting hosts.

The user traffic generation subsystem works as outlined in Fig. 8. A single generation process is supposed to be bound to each mobile host. This process generates a packet group at each arrival of a Poisson process, inserting all the packets of the group into the host's transmission queue. Since the group cardinality is distributed according to a geometric r.v., we deal with a bulking Poisson process. All packets belonging to the same group share their destination address, drawn uniformly among all mobile host's addresses. Therefore, we conclude that user traffic is offered as a sequence of point-to-point sessions, whose packets are constantly emitted, just as shown in Fig. 3. Two traffic descriptors have been chosen for the above model, the average session length (S, expressed in kbit) and the average offered load (L, expressed kbit/s). It is L = S/T, where T is the mean group interarrival time, that is the mean idle time for each generator.

Mobility events also occur at the arrivals of a Poisson process. When a mobile host enters a new subnet, it stays there for a negative exponential random time. As the time expires, an adjacent cell is uniformly drawn among the adjacent ones, and the mobile host is moved there. Since we deal with an open network, this mobility (markovian) process could be defined as a *continuous-time random walk, with discrete bidimensional states and reflecting boundaries.* The descriptor we have chosen for the mobility process is simply the average mobility rate (the inverse mean stay time).

5 Results

5.1 Analysis and Validation

An analytical estimate for the overall control traffic *induced* by MIP/ROMIP operation has been derived, according to the previously discussed model. Control traffic offered to the backbone is induced since it doesn't depend directly on the model parameters, but it is generated as a feedback of protocol and mobility events.

The following expression (dealing with average values) gives the rate R [packets/sec] at which control packets are issued by ROMIP protocol, normalized for a single user.

$$R = \frac{1}{T_{stay}} + \frac{1}{T_{stay}} + 3\frac{L}{S}\left(1 + \frac{S/B_{radio}}{T_{stay}}\right),$$

where T_{stay} [sec] is the mean stay time for a mobile host, L [kbit/s] is the mean user offered load, S [kbit] is the mean session duration and B_{radio} [kbit/s] is the available one-way bitrate on the radio channel.

The above expression includes the binding update packets released by a (idle or busy) moving host, both to its home agent (first term), and to its previous foreign agent (second term). Moreover, it includes the three control packets (warning, request, update) involved in binding acquisition for a busy source: they are issued both at each session birth (this event occurs L/S times in a second), and at each handover (during a busy period, this event occurs D/T_{stay} times in a second, where D is the session duration, that is S/B_{radio})[4].

The above expression could be recasted as follows:

$$R = \frac{2}{T_{stay}} + 3L\left(\frac{1}{S} + \frac{1}{B_{radio}T_{stay}}\right).$$

Some interesting considerations could arise by the derived formula.

- Rather obviously, the control load due to the birth of new sessions decreases by increasing the session length (at a parity of user load, L);
- The control load due to handover events could be brought down by increasing the radio channel capacity (at a parity of user load, L);
- As user load (L) increases, a proportional control load increase is induced; this reaction doesn't take place in MIP, for which is simply $R = \frac{1}{T_{stay}}$.

We have obtained a (partial) validation for our software simulator by correlating some of its output data with the above expressions. For example, the simulated system has been verified to honour the Little's formula:

$$N_{pkt} = \lambda_{pkt} \cdot T_{pkt}; \;\; \lambda_{pkt} = \frac{L}{P_{length}} + R$$

where N_{pkt} is the average number of packets in the system (user + control), normalized to a single cell; λ_{pkt} [1/sec] is the overall offered load (user + control), expressing the induced control contribution by the derived estimate (R); T_{pkt} [sec] is the mean end-to-end packet delay (obtained by weighting user and control delay); P_{length} is the IP data unit length.

For high loss values, the Little's formula is no longer applicable, even considering the system throughput instead of the whole offered traffic: in fact, a lot of packets are lost after having entered the system, thus having already affected the delay of other packets.

5.2 Numerical Results

We first analyze the mobility effects in dependence on mobility. It is useful to point out that a null mobility rate implies that all hosts remain fixed in their home networks: this means that Mobile IP protocols get excluded, and basic shortest path routing

---

[4] Leaky bucket parameters are supposed to be configured so that a single warning packet is released in the update process; moreover, packet loss is supposed to be low, so that the update process may complete without warning retransmissions.

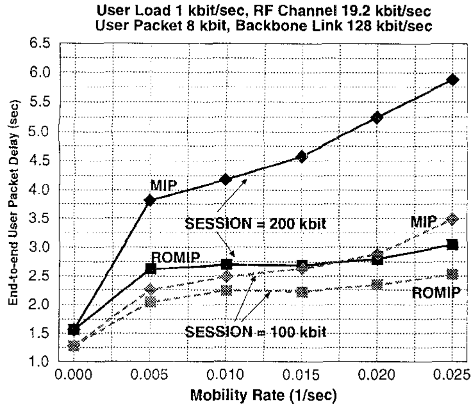

Figure 9: *MIP/ROMIP: User delay vs. Mobility (a).*

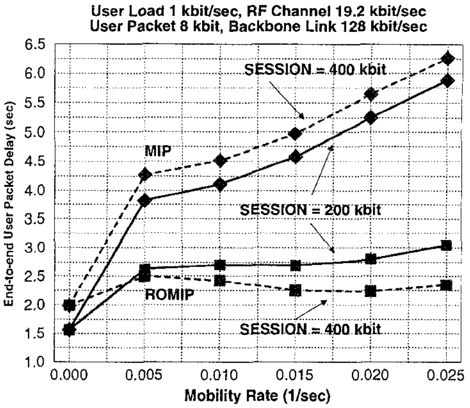

Figure 10: *MIP/ROMIP: User delay vs. Mobility (b).*

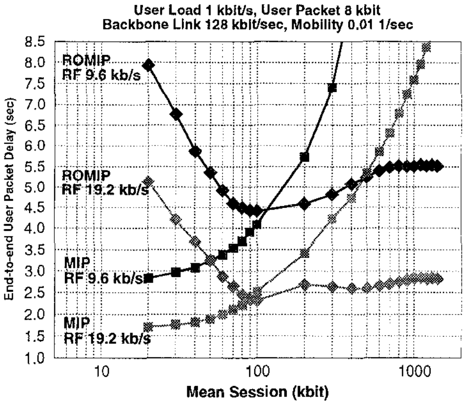

Figure 11: *MIP/ROMIP: QoS vs. Session Length.*

may be applied. Conversely, a mobility rate arbitrarily greater than zero activates redirection mechanisms: this explains the stepwise pattern observed in delay measures.

By looking at the trends shown in the diagrams 9 and 10, the following considerations could be made. Note that the average traffic offered by mobile users is always kept fixed.

- At null mobility rate, end-to-end delay always increases as session duration increases. With S = 100 kbit the minimum delay is obtained, i.e. a value slightly higher than the time needed to transmit a packet over the source and destination radio links (16/19.2 sec), meaning that any further remaining part of a delay rises up in the backbone. For null mobility, the above gap ought to be ascribed only to the increasing traffic burstiness.
- Increasing the mobility rate, the MIP delay also increases, owing to network load and tracking effort. A similar increase is observed for ROMIP too, except for the 400 kbit session: in this case we can suppose that session end-points mobility results in traffic scattering in the backbone, thus improving the delay performance over that obtained with lower mobility. With ROMIP, source and destination mobility cuts up the longest sessions into small pieces, thus canceling burstiness effects in the backbone; moreover, ROMIP may gain efficiency with longer sessions, because of the source binding acquisition process.
- Increasing session length, the MIP delay also increases, since longer and longer traffic bursts make home agents more congested (every home agent acts as a fixed routing point, even for sessions with mobile end-points). Conversely, for the values chosen in the diagrams, ROMIP seems to be much less sensible to session duration. However, it is evident from the cited diagrams that MIP delay performance improves and gets closer to ROMIP's for relatively short sessions (100 kbit).

The complex reaction revealed by ROMIP to the variations of session duration suggested to evaluate the delay performance in dependence on the session duration. The diagram 11 compares MIP and ROMIP for a broad range of session duration and for two values of radio channel capacity, giving evidence to a great difference in their QoS features.

- For short sessions (i.e. for a small ratio between session duration and mean stay time), MIP protocol achieves much lower delay than ROMIP. In fact, short sessions hardly enter their direct routing phase provided by ROMIP, because the source closes them before it can receive a valid binding for its destination. In these conditions, ROMIP degenerates and delivers packets by triangle routing; moreover, it floods the network with useless control messages, giving rise to a performance drawback.
- For longer sessions, ROMIP delay performance improves, since most of them last long enough to enter the direct routing phase. Conversely, in MIP the links surrounding the home agents rapidly become choked up by packet trains, giving rise to a huge delay (there's no flow control in our model). With ROMIP, mobility gives rise (i) to a more homogeneous traffic distribution, (ii) to a saturation effect in the traffic burstiness: so, even for the longest sessions, the time needed for delivering packets tends asymptotically to a limited value.

It is also interesting to evaluate packet loss due to transmissions to the wrong subnet (the destination is not linked to that subnet): the diagram 12 shows a slightly better performance for ROMIP, owing to its handover support. However, the difference between MIP and ROMIP loss performance doesn't appear to be substantial: ROMIP allows for in-flight packets to be recovered and forwarded, but the update process may take a long time, letting cached bindings go out-of-date. Besides, the diagram suggests that the loss probability could be reduced by increasing the backbone bandwidth, to allow a more effective tracking of mobile hosts.

As far as ROMIP overhead is concerned, we have reported some results in the diagrams 13. The right-side diagram, showing cache agent overhead for tunneling operations, depicts that a linear relation between processing load and offered traffic exists, but only for low traffic volumes. For low mobility and

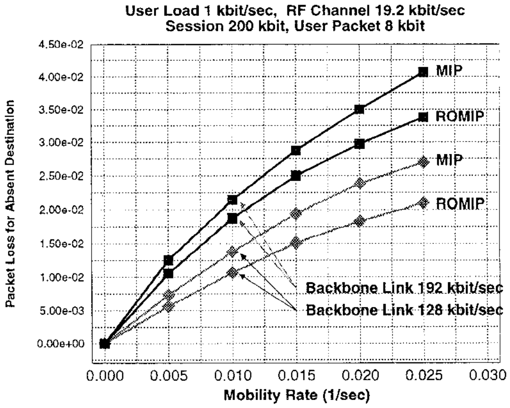

Figure 12: *MIP/ROMIP: User Loss vs. Mobility*

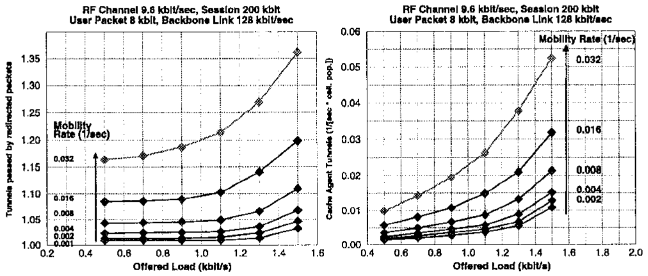

Figure 13: *ROMIP: Handover Efficiency vs. Offered Load.*

low traffic, it is revealed by the left-side diagram that the redirected packets have been tunneled only once: this is the ideal operating region for ROMIP, since a single redirection means that most cached bindings are lined up with actual locations. Mobility being equal, high traffic makes the location tracking alghoritm to lag behind: on the average, more than one tunnel hop is needed for a packet to catch the destination, binding update processes are then more frequently spurred, and so a quasi exponential relation between processing overhead and user load arises. This behaviour could seriously affect ROMIP scalability in congested environment.

We conclude our analysis by evaluating the impact of the cache size over quality of service. The binding cache tables have been supposed to be sized with a maximum number of entries equal to the average cell population. The diagram 14 explores packet loss and delay when the ratio between cache size and cell population is less than one. A first trade-off is observed: a small cache capacity gives rise to a lower loss and a higher delay, whereas a large capacity originates a higher loss and a smaller delay. In the latter context, a large amount of cached bindings may be inconsistent (this raises loss probability), but packets succeeding in reaching their destination often travel along the shortest path. To limit the packet loss due to inconsistency, our model provides for each cache entry to be removed as soon as its associated lifetime expires. A second trade-off is then observed: small lifetimes may keep the bindings up-to-date, but it is more likely that a valid binding is removed and thus triangle routing occurs, with consequent increase in the user average delay.

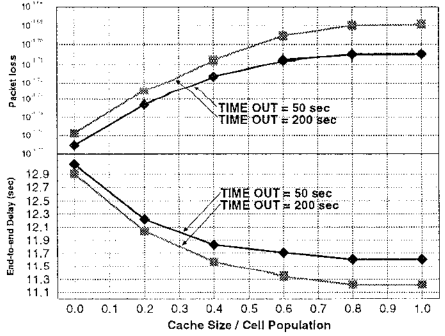

Figure 14: *ROMIP: QoS vs. Binding Cache Size.*

6 Conclusions

The performance evaluation of MIP and ROMIP carried out by simulation have led to the following conclusions:

- the indirect routing, used by MIP, shows better performance than optimized routing, used by ROMIP, as long as the rate of birth and death of sessions is high;
- large session duration better exploit the optimization of routing carried out by ROMIP;
- as long as the traffic bursts last on average as much as the average cell permanence time, the direct routing of ROMIP allows to better distribute the traffic offered to the fixed network, while indirect routing is subject to overload of the Home Agent node.

Hence, the performance behaviour of the studied protocol alternatives largely depends on the traffic pattern. Possible developments of our work are in the direction of taking into account point-to-multipoint traffic and asymmetric load distribution.

References

[1] C. Perkins (Editor). IP Mobility Support. Technical report, Internet RFC 2002, Standards Track, October 1996.

[2] D. B. Johnson and C. Perkins. Route Optimization in Mobile IP. Technical report, draft-ietf-mobileip-optim-07.txt, November 1997.

[3] C. Perkins. IP Encapsulation within IP. Technical report, Internet RFC 2003, Standards Track, October 1996.

[4] C. Perkins. Minimal Encapsulation within IP. Technical report, Internet RFC 2004, Standards Track, October 1996.

[5] J. Solomon. Applicability Statement for IP Mobility Support. Technical report, Internet RFC 2005, Standards Track, October 1996.

[6] A. Tominaga, O. Nakamura, F. Teraoka, and J. Murai. Problems and Solutions of DHCP. Proceedings of INET '95, June 1995.

[7] W. Simpson. IP in IP tunneling. Technical report, Internet RFC 1853, October 1995.

[8] S. Rajagopalan and B. R. Badrinath. An Adaptive Location Management Strategy for Mobile IP. Proceedings of MobiCom '95, ACM, November 1995.

[9] R. Yuan. An Adaptive Routing Scheme for Wireless Mobile Computing. Technical report, NEC Systems Laboratory, Inc., November 1995.

[10] R. Yavatkar and N. Bhagawat. Improving End-to-End Performance of TCP over Mobile Internetworks. Technical report, Department of Computer Science, Univerisity of Kentucky, 1995.

[11] E. W. Zegura, K. L. Calvert, and S. Bhattacharjee. How to Model an Internetwork. Proceedings of INFOCOM '96, March 1996.

[12] K. Brown and S. Singh. A Network Architecture for Mobile Computing. Proceedings of INFOCOM '96, March 1996.