PERFORMANCE EVALUATION OF TCP OVER MOBILE IP

N. A. Fikouras, K. El Malki, S. R. Cvetkovic, C. Smythe

Department of Computer Science,The University of Sheffield,
211 Portobello St.,Sheffield S1 4DP, UK

{nick, karim, srba, colin}@dcs.shef.ac.uk

Abstract-The aim of this paper is to study the effects of Triangle Routing and Mobile IP handoffs on TCP Performance. The majority of Internet hosts today do not comply with the considerations for routing optimisation. This means that until the wide deployment of next generation Internet hosts, triangle routing will be a permanent effect to Internet mobile communications. For Mobile IP handoffs it is determined that the key issue is the movement detection method (i.e. Lazy Cell Switching, Prefix-Matching, Eager Cell Switching). For this paper the performance of TCP is experimentally tested over various triangle routing lengths and through Mobile IP handoffs. For the latter all three movement detection methods are studied. The principle findings are that: triangle routing degrades the efficiency of a TCP communication by 1% per hop in the network path. TCP communications over Mobile IP handoffs can be brought to a halt for 5.5 seconds and 12.5 seconds when based on the Eager and Lazy Cell Switching methods respectively. The network scenario considered involves single-agent subnetworks and mobile nodes that do not have the capability to contemporarily participate in multiple networks (e.g. IEEE802.11).

I. INTRODUCTION

A. The Mobile Internet Protocol

TCP/IP was originally designed without any considerations for mobile computing. Its greatest resistance to mobile computing resides in the routing service of the IP layer. Mobile IP (MIP) [1]] is an enhancement to the IP protocol which allows transparent routing of traffic to mobile nodes in the Internet. MIP introduces three new entities required to support the protocol: the Home Agent (HA), the Foreign Agent (FA) and the Mobile Node (MN) but may alternatively operate without FAs.

According to MIP, the mobile node is required to notify the Home Agent of its current network access point. For every supported MN the Home Agent creates a record (registration) of its current location and acts as its proxy into the Home Network (HN). Intercepted incoming traffic is redirected by the Home Agent to the mobile node's registered location. Mobile node traffic to other nodes bypasses the Home Agent and is delivered directly. The visualisation of the aforementioned traffic flow resembles a triangle and is therefore referred to as triangle routing (Figure 1).

The HA redirects traffic through encapsulation. There are several encapsulation schemes but in this paper only IP in IP encapsulation [2] is assumed. Furthermore, it is assumed that the overhead of the encapsulation process is insignificant, as it only increases the size of an IP packet by 20 octets.

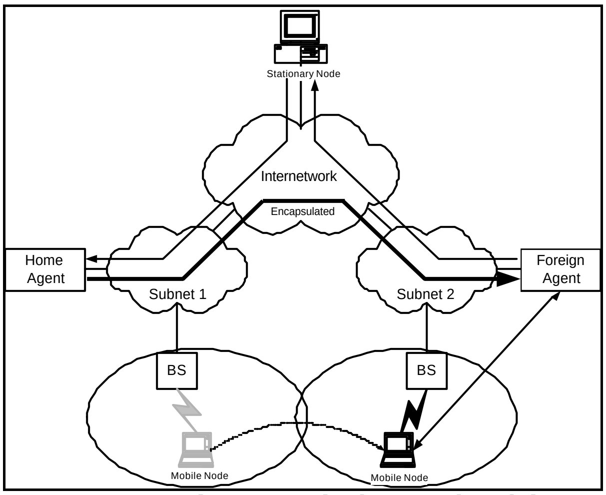

Figure 1: Two departmental subnets with mobility support

A typical scenario is presented in Figure 1. It illustrates two departmental Internetworks (subnets), called Subnet 1 and Subnet 2 which may consist of various LAN technologies including Wireless LANs (WLANs). The boundaries of the two subnets meet at their WLANs where they overlap. This area from now on will be referred to as the overlap area. A Mobile Node (MN) that has been permanently allocated an IP address into Subnet 1 moves through the overlap area into Subnet 2. Assuming that Subnet 2 does not offer mobility support (normal IP), the MN will find that it cannot access the network. Moreover, other nodes attempting to communicate with the MN will find that it is unreachable. In these cases, MIP can provide transparent routing of MN traffic to its new location in Subnet 2.

Node Mobility has inherent the act of MN handoffs. In general, a MN that leaves a network (LAN or IP subnet) and enters another must perform a handoff. Its purpose is to re-associate the MN with a designated entity that offers network connectivity in the new network. In Internetworks with mobility support such as the one presented in Figure 1 two different types of handoffs take place. The first is at the link (OSI layer 2) layer and the second at the IP (OSI layer 3) layer that respectively maintain link connectivity and network access.

MIP introduces the overheads associated with triangle routing and IP layer handoffs (MIP handoffs). In this paper the impact of these overheads to TCP performance is determined. In particular, the degradation of TCP performance per hop in the triangle route is experimentally obtained. Furthermore, it is shown that the TCP mechanisms of timeout exponential backoff [3]and Path MTU Discovery [4, 5] cause serious problems to a communication when mobility is introduced.

II. BACKGROUND MATERIAL

A. Mobile node handoffs

In section 1 two distinct types of handoffs were identified that take place at the link and the IP layers. Their aim is to provide link connectivity and network access respectively. The term link connectivity indicates that the MN has a physical connection to a LAN. In WLANs with roaming support a link layer handoff is the process by which a MN creates an association with a Base Station (BS) that provides the link connectivity. In this paper we assume that link layer handoffs complete seamlessly and almost instantaneously. The term network access denotes a MN's capability to exchange traffic in its current subnet without compromising its permanently allocated identity (IP address). A MN is said to have network access when its current subnet location corresponds with its registered network access point. The latter is the location indicated by the MN's HA registration. Therefore, a MN whose location does not correspond to its registered network access point will not have network access even though it has link connectivity. In the absence of network access capability it is considered that the MN is experiencing service disruption.

Service disruption terminates when the HA registration is updated by a MIP handoff. From that it is derived that in the absence of an overlap area the disruption interval is equivalent to the duration of the MIP handoff. In the presence of an overlap area it is up to the LAN technology to avoid this disruption by enabling the MN to receive traffic from both networks that form the overlap. In this case the MN that is located between two subnets will be in the position to use its registered network access point for network access while it is initiating a MIP handoff to update the HA registration. If the MN stays in the overlap area longer than the MIP handoff interval, no service disruption will be experienced. The capability of a LAN technology to deliver traffic from multiple networks while in their overlap area is defined as "true overlapping". An example of a technology that does not offer true overlapping is the IEEE 802.11 standard for Wireless LANs [6]. In this paper it is assumed that the LAN technology does not support true overlapping and therefore the service disruption interval experienced by a MN during every location switch is equivalent to the duration of the MIP handoff.

B. Mobile IP movement detection methods

The only mean by which a MN's is made aware of its location is through evaluation of periodically received agent advertisements that are broadcasted by agents. Every agent advertisement contains a number of advertised network access points. The receipt of agent advertisements that advertise the same access point as the MN's registered network access point indicates that the MN has not moved. Loss of contact with that particular agent for 3 successive advertisement periods could indicate that the MN has relocated outside the agent's subnet. Furthermore, the receipt of advertisements from a new agent in a new subnet could be perceived as indication of movement into a new subnet. In either of the last two cases the MN might decide to handoff, remain as is or give up a network access point. The decision is up to the movement detection method that the MN deploys. In this paper the three known movement detection methods will be considered, namely: Lazy Cell Switching (LCS) [1], Prefix Matching (PM) [1] and Eager Cell Switching (ECS) [7].

The LCS method has the main characteristic that after the MN has registered a network access point with the HA it remains devoted to it and ignores any newly discovered ones. This is followed until the MN misses 3 successive agent advertisements indicating that it has left the subnet. In that case the MN must discover and register a new network access point. It be will shown in following sections that the LCS has the longest movement detection interval and therefore claims the biggest impact to TCP [8] performance.

The PM method has a similar functionality to the LCS with the only difference being that it includes a prefix-length extension to agent advertisements. This extension contains the prefix length of the agent's IP address and it enables the MN to determine whether the agent advertisement has been received from an agent in the same or a new subnet. The later would indicate that the MN is closing into another subnet and perhaps it should initiate a handoff to avoid service disruption. This case is

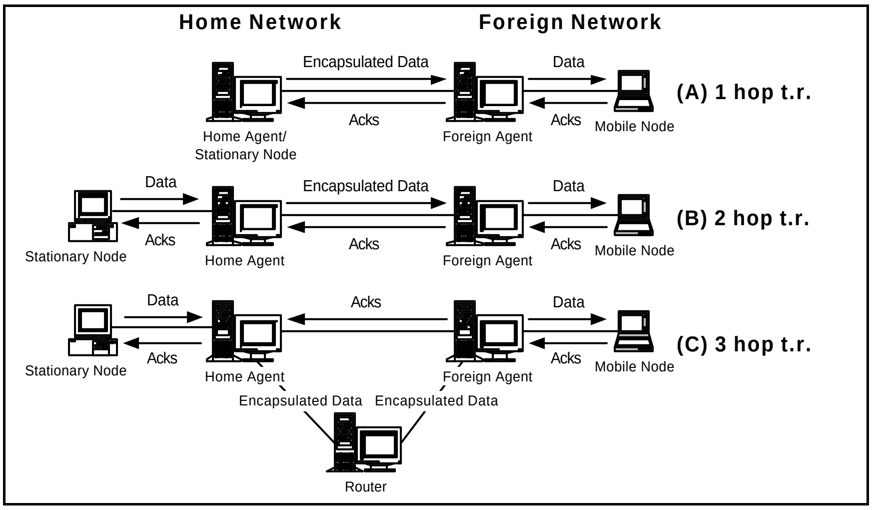

Figure 2: A TCP communication with (A) 1 hop (B) 2 hops (C) 3 hops triangle routing.

only possible when the LAN technology supports true overlapping. In the absence of this capability PM tends to operate like the LCS and therefore a dedicated study of TCP performance over MIP handoffs with PM support has been omitted from this paper.

The ECS method tends to function in a way opposite to that of the LCS. It assumes that MNs change their direction of movement very slowly. That is, if they are moving forward in one direction it is unlikely that they will stop and turn back. Hence, it is appropriate for MNs to handoff immediately upon discovering a new access point. This method tends to reduce the movement detection interval in comparison to the LCS method and therefore manages faster MIP handoffs. The study of MIP handoffs with ECS support presented in this paper assumes only single-agent subnets. Research on the behaviour of the ECS in multi-agent subnets is under way.

III. EXPERIMENTAL SETUP

This paper analyses the performance of TCP over MIP. It focuses on the effect of the triangle routing and MIP handoff overheads to TCP performance. For the latter only MIP handoffs with LCS and ECS support are studied because it is assumed that in the absence of true overlapping the PM method functions the same as the LCS method. For each test series, either triangle routing or MIP handoffs a different MIP implementation is used. The first test series is performed over version 2.0 beta of the National University of Singapore (NUS) Mobile IP implementation [9] for the Linux Kernel 2.0.24 that is compliant with [1]. For the second series the Sunlabs Mobile IP [10] implementation for Redhat 5.x Linux is used. It is fully compliant with [1] and [2] and implements all three movement detection methods. Figures 3 and 4 present the testbed topologies for each test series. They consist of Pentium PCs that are running RedHat 5.2 Linux. All network connection is provided with Ethernet. The tool used for the generation and benchmarking of TCP traffic is the public-domain *ttcp* benchmarking software [11].

IV. BENCHMARK RESULTS

A. The triangle routing overhead

The current infrastructure of Intenet hosts does not comply with the considerations for routing optimisation to mobile communications that eliminates the need for triangle routing. Consequently, until the wide deployment of next generation Internet hosts, triangle routing will be a permanent effect to all Internet mobile commuincations.

It is well known that extending the path of a communication by additional hops degrades performance. The triangle routing test series do not aim in pointing this out, rather to raise awareness on the impact of triangle routing to TCP performance. In this section the amount by which TCP performance is degraded per single hop in the triangle route is identified. For this, three different tests have been performed in which the triangle route is increased by a single hop. TCP communication have been tested with: (A) 1-hop triangle route, (B) 2-hops triangle route and (C) 3-hops triangle route. In each test 1000 MB is transmitted over a “raw” TCP connection from a Stationary Node (SN) to a MN. The main characteristic of this transfer is asymetry to the amount of traffic that the MN sends and receives. That is, the MN is the receipient of a bulk of data while it is required to respond with small acknowledgements. Only communications of this type are considered because it is expected that MNs will mostly participate in such communications (i.e. Web traffic). Each transmission is repeated for ten times to eliminate randomisation. The experimental topology corresponding to each test is presented in Figure 2. In all three tests the link between the HA and the FA is provided with two Ehernet links to eliminate Carrier Sense Multiple Access/Collision Detection (CSMA/CD) collisions and preserve resemblance to real life scenarios.

The results of the three test cases are presented in Table 1, in terms of communication efficiency. It is defined as:

$$Efficiency = \frac{file\ size(MB) \big/ transmission\ delay(seconds)}{maximum\ capacity\ of\ the\ Ethernet\ network(MB/sec)} \times 100$$

The results indicate that for each hop added to the triangle route, TCP normalised throughput performance drops almost 1%. Assuming that in a typical TCP communication over MIP a node would experience a triangle route of 10-15 hops every TCP communication of this type should be prepered to face a 10% to 15% drop to TCP efficiency. From the tests it is identified that a 20% drop to efficiency corresponds to a 50% increase to the TCP communication delay. The level of decrease is too small to define an upper bound limit to the length of the triangle route but it is considered that any further decrease to TCP efficiency beyond 20% will seriously affect user satisfaction.

Another TCP mechanism with resistance to mobile computing is the Path Maximum Transmission Unit (MTU) Discovery. The functionality of this mechanism is to drop packets that are larger than the Path MTU. In these test series it is identified that MIP suffers from this mechanism as MN movement is usually accompanied either by encapsulation packet size increase or Path MTU reduction due to change in the network path. In either case Path MTU Discovery causes loss of at least one TCP packet which can only be recovered through a TCP timeout. It is considered that due to Path MTU Discovery a TCP communication over MIP must suffer an extra TCP timeout with every MIP handoff. It is noted that TCP communications that suffer timeouts must commence the communication with the slow-start algorithm [12, 13].

Single hop triangle route	76.56
Double hop triangle route	75.63
Triple hop triangle route	74.67

Table 1: The efficiency of a TCP communication with an increasing triangle route

In the MIP handoff tests series two cases of MIP handoffs with LCS and ECS support respectively, are presented. The testbed topology for these tests is presented in Figure 3. In every experiment the transmission rate of agent advertisements is set to the recommended by [1] value of one per second. A MN will consider a network access point invalid after missing three successive agent advertisements. This means that MN's will give up an access point after lossing connectivity with the subnet agent for 3 seconds.

In Figure 4, it can be observed the Path MTU Discovery mechanism causes the loss of the first transmitted packet after the completion of the MIP handoff. It can also be observed that the TCP timeout interval is doubled with every successive timeout. From Figure 4 it is derived that ECS based MIP handoffs complete 3 seconds faster than their LCS counterparts. This enables the first TCP to recover 7 seconds earlier than the latter one. In both recoveries TCP traffic commences with the slow-start algorithm. It is considered that its effect is insignificant compared to that of TCP exponential backoff.

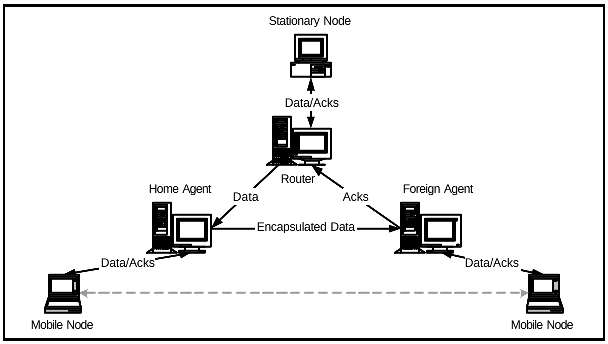

Figure 3: For the Mobile IP handoff tests the mobile node switches its location while sustaining a TCP communication with the stationary node

V. CONCLUSIONS

In this paper the performance of TCP over Mobile IP has been evaluated. The study focuses on the impact of triangle routing and Mobile IP handoffs to TCP performance. It is identified that triangle routing degrades the efficiency of a TCP communication by 1% per intermediate hop in the triangle route. It is presented that a 20-hop triangle route increases by 50% the delay of a TCP communication. The study of Mobile IP handoffs assumes single-agent subnetworks and inability of mobile nodes to contemporarily participate in multiple subnetnetworks. It is identified that the duration of a Mobile IP handoff is directly dependent on the movement detection method. Under the assumed conditions it is presented that the Prefix-Matching method has a similar functionality to the Lazy Cell Switching, therefore only the Lazy and the Eager Cell Switching methods have been considered. It was experimentally derived that Eager Cell Switching based Mobile IP handoffs complete 3 seconds faster than their Lazy Cell Switching counterparts. It was also shown that during a MIP handoff TCP experiences several successive timeouts that increase the timeout interval beyond the duration of the handoff. This causes the TCP communication to remain halted even after the completion of the MIP handoff. It was also identified

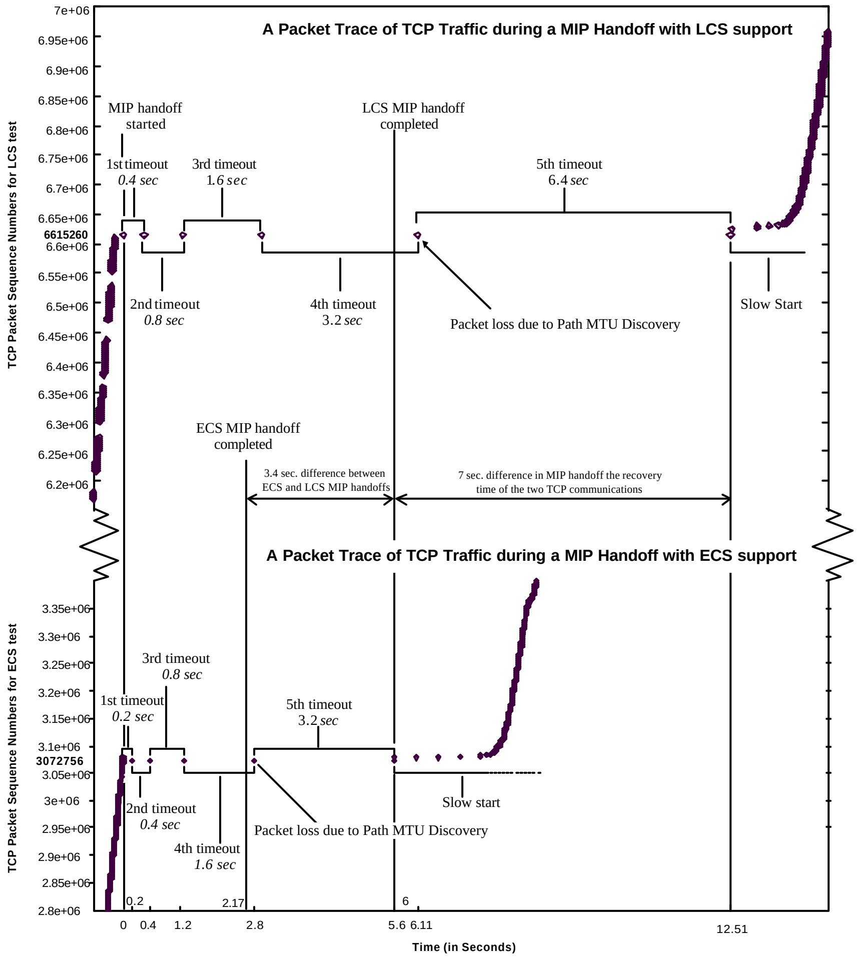

Figure 4: A comparison of TCP packet traces during Mobile IP handoffs with LCS and ECS support

that movement is usually accompanied either by packet size increase or Path MTU reduction which leads to further packet loss by the Path MTU Discovery mechanism. In total it is experimentally identified that TCP recovery of a Mobile IP handoff with Lazy Cell Switching support is 12.5 seconds while the corresponding delay for Eager Cell Switching handoffs is 5.5 seconds.

VI. REFERENCES

[1] Perkins C., "*IP Mobility Support.*", RFC 2002, (Oct. 1996).

[2] Perkins C., "*IP Encapsulation within IP.*", RFC 2003, (Oct 1996).

[3] Stevens, R., "*TCP/IP Illustrated Volume 1.*", Reading: Addison Wesley Professional Computing Series, (1995).

[4] Kent, C. and Mogul, J., "*Fragmentation Considered Harmful.*", In Proc. SIGCOMM '87 Workshop on Frontiers in Computer Communications Technology, (Aug. 1987).

[5] Mogul, J. and Deering, E., "*Path MTU Discovery.*", RFC 1191, (Nov. 1990).

[6] IEEE Std 802.11-1997. "*Wireless LAN Medium Access control (MAC) and Physical Layer (PHY) specifications.*", (1997)

[7] Perkins C., "*Mobile IP, Design Principles and Practices.*", Addison Wesley, (1998).

[8] Postel B., "*Transmission Control Protocol.*", RFC 793, (Sept 1981).

[9] NUS 2.0beta Mobile IP, [ftp://ftp.nus.sg/pub/mobile-ip/]

[10] Sunlabs Mobile IP, [ftp://playground.sun.com/pub/mobile-ip/]

[11] "*ttcp: Test TCP.*" U. S. Army Ballistics Research Lab, (Dec 1984).

[12] Jacobson, V. and Karels, M. "*Congestion Avoidance and Control.*", (1988), [http://www-nrg.ee.lbl.gov/papers/congavoid.pdf]

[13] Stevens R., "*TCP Slow Start, Congestion Avoidance, Fast Recovery Algorithms.*", RFC 2001, (Jan 1997).