

Performance Evaluation of Transport Protocols with Local Mobility Management

Nikos Georganopoulos* and A. Hamid Aghvami

Centre for Telecommunications Research,
Kings College London, Strand,
London WC2R 2LS, UK
n.g@ieee.org
hamid.aghvami@kcl.ac.uk

Abstract. There is an increasing interest to design a full IP Access Network that can provide IP connectivity to mobile users. UDP and TCP transport layer protocols are designed to deliver data across wired networks. Mobility of users in such a wireless system results in transport layer protocols facing a very different packet loss and delay pattern. The use of an efficient mobility management protocol is essential to the performance of the protocols and the relevant applications running on top. In this paper¹ the performance of UDP and TCP transport layers is evaluated on a wireless IP access network running the BCMP localised mobility management protocol, through simulations using the ns-2 simulation platform. Data and multimedia applications are employed to test their performance. UDP and various TCP variants are tested and the objective is to evaluate their performance in such a system and to examine the effect of the mobility of users.

1 Introduction

There are currently numerous new data applications being developed and provided in the Internet. At the same time important applications like voice and video are to be offered also over IP with the use of developed QoS mechanisms. Furthermore, the philosophy behind the IP protocol stack is regarded to be better than that of traditional networks such as ISDN or ATM that are connection oriented employing cell or frame switching. Finally the growth of the Internet is rapid and is extending in terms of data traffic volumes, users and geographical locations. All these reasons prompt the use of IP for a wireless access network that will provide a vast number of services to its users, on the move. As a result there is currently ongoing research effort to design such a system that can provide IP services and applications to mobile users through a mobile-fixed converged backbone network. These systems could be considered as next generation cellular system offering full IP connectivity to its users.

* Currently with: Toshiba Research Europe Ltd, 32 Queen Square, Bristol BS1 4ND, UK.

¹ Part of this work has been performed in the context of the IST consortium BRAIN (IST-1999-100 50) and MIND (IST-2000- 28584), which are partly funded by the European Commission. The authors would like to thank their colleagues for their contribution in this work.

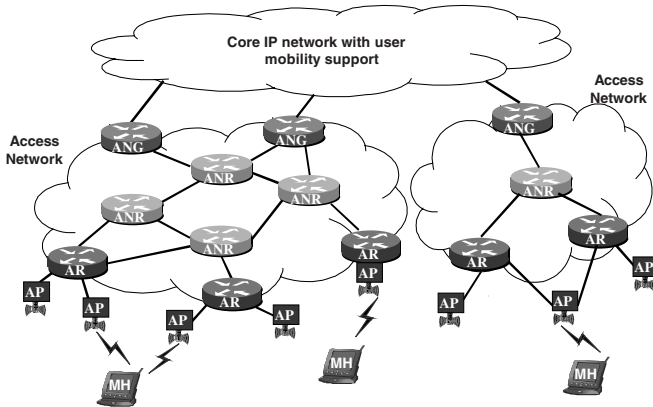


Fig. 1. Access Network Reference Architecture, including one or more Access Network Routers (ANR) and is connected to the Core Network (CN), a ‘standard’ IP network, through an Access Network Gateway (ANG). Access Router (AR) are special ANRs residing on the edge of an AN and connected to one or more Access Points (AP), a layer-2 device that is connected to one or more ARs and offers the wireless link connection to the Mobile Host (MH). An AR offers IP connectivity to MHs, acting as a default router to the MHs it is currently serving

1.1 Access Network-Local Mobility Management Domain

There is thus a strong motivation behind the design of an IP wireless access network system. A high level description of such a network topology can be found in [1] where definitions of the different system entities are given. Such a system will manage mobility of the users locally inside its administrative domain through the use of a Localised Mobility Management (LLM) [2] solution. Figure 1 shows a general diagram of such wireless access network architecture. The main network components in the architecture are the following [1]. The Access Network (AN) (Fig. 1) is a full IP network.

An LMM solution will be running inside the AN on specific ANRs that will be responsible for tracking the movement of the MHs at the edge of the network and ensuring that packets are timely delivered to every MHs’ current point of attachment to the network.

1.2 Transport Layer Protocols

User Datagram Protocol (UDP) is a connectionless transport protocol that provides the fastest way for applications to send messages and is used by time sensitive multimedia applications. Transmission Control Protocol (TCP) provides a reliable data-transfer and is used for bulk data traffic and interactive data applications. TCP is the mostly used transfer protocol and supports the transfer of over 90% of all traffic in the public Internet today. Different variants of TCP have been proposed and are considered in this study. These include Tahoe, Reno, Vegas, NewReno, Sack and Fack. Support of wired Internet will be inherit in such a system, thus rendering

necessary the evaluation of the performance of transport layer protocols and the effect of the mobility of users

2 Mobility in the Wireless IP Access Network

In theory, transport protocols should be independent of the technology of the underlying layers and particular the transmission media at the physical layer. However the wireless IP-based AN will be designed to provide IP services and applications to MHs, over the air where user mobility is of primary concern. There are thus two main characteristics of the AN that will have an effect on the performance of the transport protocols:

- **Host (terminal) mobility:** Packets can get lost as hosts move and handover between different ARs residing at the edge of the network. The path that packets follow in the AN towards the MHs changes and is updated to track their movement. Packets can then be lost or reordered during this update and the applications will suffer. The impact of host mobility on the transport protocol performance is mainly studied and presented in this paper, together with the interaction with the chosen mobility management solution that is running in the AN.

- **Wireless channel:** Packets will be lost because of bad performance of the wireless channel (fast and slow fading). This will have an impact on applications running on top of the transport protocols. The architecture of the network considered in this study is independent of the chosen air-interface, which is though dependent on the actual system implementation. This aspect thus is not considered in this study. There are various other proposals to improve the performance of transport protocols over wireless links.

2.1 A Localised Mobility Management Solution

The mechanism used to track the movement of the MHs at the edges of the AN is called LMM protocol. To manage handovers between domains a global mobility management protocol is needed like Mobile IP (MIP) [3]. An LMM protocol is needed to provide either fast (minimal packet delays) or smooth (minimal loss) handovers. Furthermore, local mobility should be managed as transparently as possible to the nodes outside of the access network. These solutions include protocols like Hierarchical MIP, Cellular IP, HAWAII.

In this study a new LMM proposal, called BRAIN Candidate Mobility Management Protocol (BCMP) [4] is employed. Only limited number of nodes need be modified to support BCMP. The network consists of legacy IP routers with added mobility aware functionality in just two types of nodes. Anchor Points (ANPs) own and allocate IP addresses, authenticate users, maintain records, and tunnel packets towards MHs. ARs terminate tunnels from ANPs and forward packets to/from mobile hosts. By supporting the ability for the ANG to change, the BCMP provides a resilient and scalable solution. The main protocol procedures dealing with mobility are:

- **Handover and Path Updates:** This function includes an optional handover preparation phase to ensure fast and smooth change of ARs by performing a planned handover. The handover execution procedure is the same whether or not there has

been a preparation phase (unplanned handover). The ANP is also notified about the new location of the MH and the path is updated.

- **Inter-Anchor Handover:** With this function the protocol allows (but does not mandate) the network operator to request that a MH changes ANP if it has moved far away from its original ANP. This improves routing efficiency in the AN in exchange for exposing mobility toward the Internet.

Furthermore login, logout and paging procedures are defined but are not relevant to the study presented here.

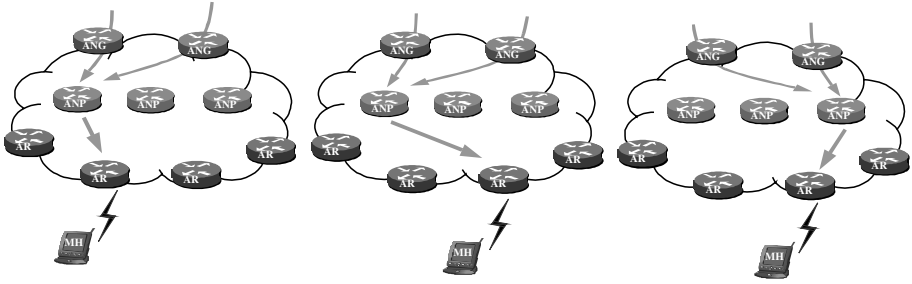


Fig. 2. BCMP Operation

3 Experimental Set-Up

The evaluation of the performance of transport layer protocols has been performed using the ns-2 simulation platform [5]. As mentioned above different network topologies are employed. These include a tree topology and a mesh topology. Regarding BCMP [4] has shown the best configurations of the protocol entities to get the best performance in terms of handover delay and application performance. This configuration requires to put the ANPs in the middle of the AN, that is nodes 10, 11, 12 in the tree topology and nodes 14, 16,18 in the mesh topology.

Nodes 0 to 5 represent various CHs and node 6 in this study represents the MHs' HA that forwards the packets to the current address of the MH. One of the properties of the BCMP is that it is independent of the global mobility protocol but for this study the use of MIP is assumed. Although there is no mobility between different ANs, the MH can change its assigned address inside the AN when changing ANPs something that will trigger a global mobility event. This also depends on the actual size and physical topology of the AN. Furthermore the protocol supports multiple ANGs, which in the mesh topology are in nodes 7 and 9. In the tree topology there is only one ANG in node 7. A static routing protocol is employed in the AN. Regarding the physical characteristics of the network, the following table summarises the different values used. These values are the same for both network topologies.

There are 24 MHs attached to the AN, only 2 of which will be moving in the AN, performing handovers. Moving MHs will be moving at the edge of the AN performing a total of 100 handovers every 1 second. It is assumed that cell coverage of the areas of the ARs is overlapping, thus ensuring a graceful physical handover from one AR to the other. The wireless channel is assumed to be perfect, with no packet loss occurring.

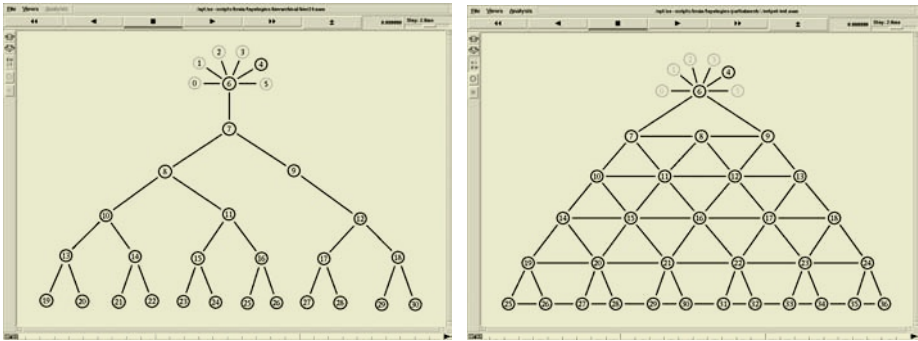


Fig. 3. Tree and Mesh Topologies

Table 1. Network Physical Characteristics

Link	Bandwidth	Delay
CH(s) – HA	10 Mbytes	5 ms
HA – ANG(s)	100 Mbytes	5 ms
AN wired links	20 Mbytes	1 ms
AN wireless links	10 Mbytes	0.5 ms

Various applications will be supported by the MHs, which will produce a mixture of realistic traffic throughout the duration of the simulations. These applications will be running over UDP and TCP. This will enable to study the effect of mobility on the transport protocols.

- **UDP Applications:** VoIP is one application implemented, which is a two-way communication between the MH and the CH. Streaming video is another application that is used over (RTP over) UDP, to transmit a movie from a Video Server to the MH. The employed video model is a trace-driven one, using a real movie trace file (Jurassic Park I). Finally a videoconference application (over RTP over UDP) is also simulated and used to generate traffic load. For such an application real video traces are used again.

- **TCP Applications:** Web browsing is one application that relies on TCP to transfer information from a web server to the MH. For this application, a web-browsing model, included in the libraries of ns, is used. File transfer with FTP is the main TCP application that is used during the simulation. This application is run throughout the duration of the simulation, to achieve a maximum file size transfer at the end of the simulations. The achieved throughput is controlled by TCP, as it will be illustrated later in the document. The different variables of the TCP connections can also be set for the simulation. The packet size of the TCP packets is 1480 bytes, and the maximum window is varied to plot graphs with varied advertised window size.

4 Performance Evaluation

This section will describe the results of the performed simulations. Results for UDP and TCP will be given in two different sections starting with UDP.

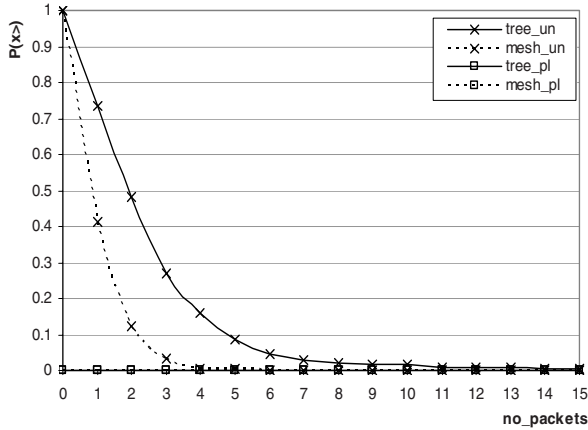


Fig. 4. Packet Loss Probability per Handover

Analysis of Experimental Results for UDP: The graph of Fig. 4 plots the probability of packet loss per handover because user mobility. It can be seen that when performing planned handovers user mobility can be completely hidden, as there is zero packet loss in both topologies. In the case of unplanned handovers packets can be lost and in the case of tree topology there is a high probability of even up to 5 packets been lost during a handover going up to even 14 lost packets. In this case the topology has a big impact on the number of lost packets since in the tree topology, for specific handover the handover messages have to go through the ANG before reaching the old or new AR (handover between nodes 26 and 27 in Fig. 3). In the case of the mess topology only 1 or 2 packets have a high loss probability with 5 the higher value that can occur.

One metric is the number of packets with higher than a specific end-to-end delay variance value, where delay variance is the difference between the time the packet arrives at the destination and the time the packet was generated at the sender plus the time it takes to forward the packet through the network links end-to-end. This is similar to IPPM’s One-Way Delay Metric defined in [7] minus the end-to-end forwarding delay, including transmission and propagation delay, thus leaving the queuing and mobility induced delays. This metric thus shows the number of packets that will be lost depending on the playout delay offset of the specific application and its decoding process. In the mesh topology case results show that the end-to-end delay variance is increased by around 0.5 ms for up to 5% of the packets with smaller values for planned than unplanned handovers. In the tree topology, up to 10% of the packets face a higher end-to-end delay variance even up to 4 ms, with higher values for planned than unplanned handovers. Actual graphs are not included due to limited space. Note that the forwarding delay through the network is 15-16 ms, resulting in an end-to-end delay variance ranging from 10% to 40% of that delay.

Next graphs in Fig. 5 show the probability of the delay jitter, which is the relative difference between the arrival of two consecutive packets in sequence. This is similar to IPPM’s IP Packet Delay Variation Metric described in [8] and assuming that the packets are in sequence.

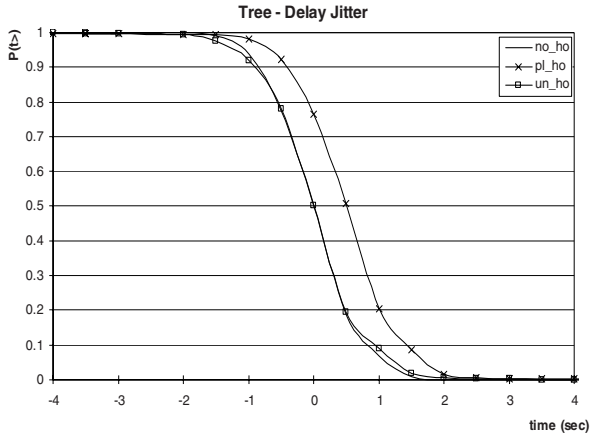


Fig. 5. Packet Delay Jitter Probability

This metric is important for some multimedia applications as they use this metric in the decoding process and ignore packets if they are late compared to the previous packet in the sequence. This metric can take negative values as the difference between the inter-arrival time between two packets in sequence can be smaller than the difference between their relevant generation time. However, what is important is the positive side, since packets with a value of delay jitter higher than the delay offset of the application will be ignored. The experiment showed that in the mesh topology mobility of the users does not have an effect on the delay jitter as the three plots are almost the same. However in the tree topology when planned handovers are performed, packets have a higher probability of larger delay jitter. This is because of the fact that packets during planned handovers have to be forwarded from the current AR to the new AR, so that they don't get lost and this redirection can include various more links thus increasing the jitter delay.

Analysis of Experimental Results for TCP: When evaluating the performance of TCP and since it guarantees delivery of all packets the major metric is the achieved throughput of the connections. Although the performance of the TCP variants has been compared before and this comparison has been the base for designing more variants, this has happened considering only the problem of congestion in the network. However considering mobility modifies the problem since the network is not congested and the packet loss pattern is very different. Thus, the performance evaluation performed here will show the effect of mobility on TCP and its variants.

The experiment showed that without mobility, with this network configuration, up to a window size value of 30 the TCP variants perform identical since congestion doesn't occur anywhere in the network. For higher max window values there is a small differentiation because of congestion with all the TCP variants achieving 98% to 100% of the 10Mbps throughput, which is the bandwidth of the CH to HA link. When mobility is included and planned handovers are performed the results are similar, showing that again planning a change of AR can hide the mobility of the users and eliminate packet loss.

However in the mesh topology (Fig 6) and in the case of unplanned handover the performance of the various TCP variants is differentiated since packets are lost during

handover (up to 30 max window) and furthermore because of congestion (more than 30 max window). Fack and Sack show the best performance followed closely by NewReno and Tahoe that outperforms Vegas and Reno.

In the tree topology (Fig. 7) the handover delay is larger and thus more packets are potentially lost, which results in a higher differentiation in the performance of the TCP flavours. Like in the mesh topology Fack and Sack show the best performance achieving 98%, 99% of the maximum throughput. NewReno and Tahoe show a very good performance of 95% achievable throughput whereas Reno and Vegas suffer achieving only 65% to 70% percent of the achievable throughput.

Assuming that packet loss because of mobility and congestion is the main reason, the performance of TCP still depends on various further factors like how many consecutive packets were lost, or where in the whole window did this packet loss occur. Consequently, is hard to come to absolute conclusions when evaluating the performance of TCP and its variants. However some firm conclusions can be well supported and more available results are not included here because of limited space.

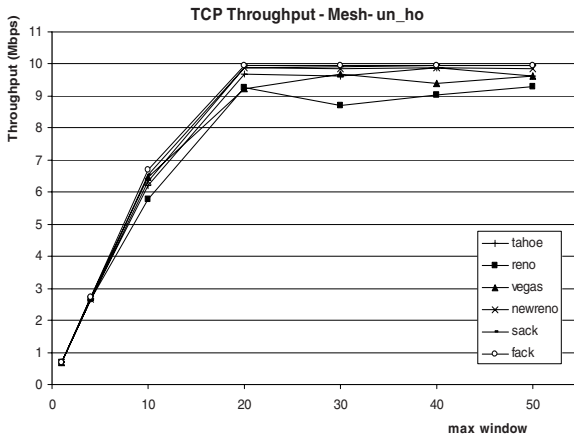


Fig. 6. TCP Throughput – Mesh Topology

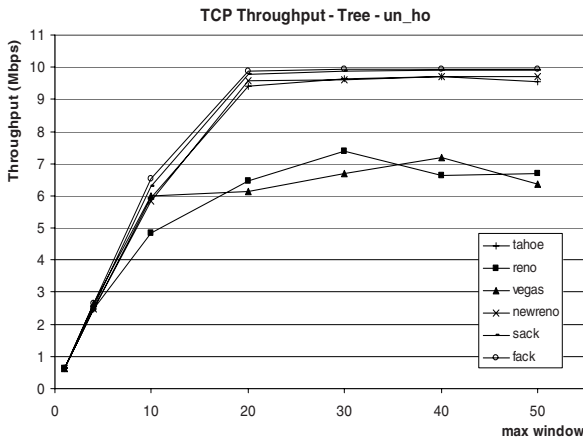


Fig. 7. TCP Throughput – Tree Topology

5 Conclusions and Future Work

This article has presented the results of the study on the performance of standard transport layer protocols over a wireless IP access domain with host mobility locally managed with the use of BCMP. A software experimental platform has been built based on ns-2, using standard data and real-time multimedia applications running on MHs moving inside the access network and performing handovers between ARs. The topology of the access network has also been varied including mesh and tree topology. IP transport layer including UDP and TCP with its various flavours have been employed.

The focus of the study was on the effect of the mobility of the users inside the access network. It has been shown that mobility may have an effect on the performance of the protocols. Detailed graphs were not included due to limited space. When trying to evaluate these results the following conclusions can be drawn:

- Successful planned handovers can hide mobility entirely from UDP and TCP. Packet loss is eliminated, not affecting UDP applications and aiding TCP to achieve maximum throughput, irrespective of the variant used. This is irrespective of the topology of the network, as mesh and tree topologies show the same performance. End to end delay variance is minimised and slightly increased packet delay jitter can only occur in the tree topology.

- In terms of network topology, mesh topology shows an excellent performance from all aspects of the system for both UDP and TCP. Of course mesh topology involves more nodes and links making it more expensive. Furthermore more efficient routing protocols (QoS routing) can be used to further improve the performance, thus increasing the complexity of the system. Tree topology, which is smaller and simpler, also shows a very good performance especially when planned handovers can be successfully executed by the MHs.

- When comparing the various TCP variants, interesting results can be drawn when studying the case of packet loss occurring because of unplanned handovers only or together with congestion. **Sack** and **Fack** with the use of selective ACKs yield the best performance, which is expected since these techniques use a different type of ACK that can acknowledge reception of multiple packets in a window. Bearing in mind that in this mobility scenario multiple packet loss per window is very probable, very good performance is achieved, yet the complexity of the algorithm and the size of the ACK packets are increased. **Reno**, which is currently widely used in the Internet, yields the worst performance, around 65% of the maximum possible throughput, together with **Vegas** that also yields a very poor performance. This is justified since Reno is often forced to wait for a retransmission timeout to recover from two or more packets dropped from the same window of data [9], which is what happening in this scenario. Regarding Vegas, although its flow control mechanism is very well designed and no congestion ever occurs, it suffers when trying to recover from packets that are lost because of handovers. **NewReno** also exhibits very good performance from the family of TCPs that uses a standard receiver, i.e. one that sends a single ACK per received packet. So NewReno with small modifications to the basic Reno implementation does give a big improvement on its performance. Finally **Tahoe**, which is the basic implementation without Fast Recovery, yields a very good overall performance, outperforming Reno and Vegas. This could be justified as when facing multiple packet loss regularly (every handover) Tahoe TCP by halving its

widow size and re-transmitting all the packets starting from the first lost, manages to transmit more packets in the long run of the experiment. Tahoe has shown similar good performance in a similar study involving a different localised mobility protocol in [10].

More detailed analysis and graphs could justify the results presented here. Further, work in this area should include the use of a realistic wireless channel and the study of further TCP versions designed for wireless systems like M-TCP, Snoop protocol, I-TCP and others. Finally, layer-4-aware layer-2 protocols can be developed to improve the performance TCP and UDP over wireless links.

References

1. J. Manner, M. Kojo, C. Perkins, T. Suihko, P. Eardley, D. Wisely, R. Hancock, N. Georganopoulos, "Mobility Related Terminology" IETF draft, draft-ietf-seamoby-mobility-terminology-04.txt, (work in progress), April 2003.
2. C. Williams, "Localized Mobility Management Requirements" IETF draft, draft-williams-mobileip-lmm-requirements.txt, (work in progress), July 2001.
3. C. Perkins, "IP Mobility Support for IPv4", IETF RFC 3220, January 2002.
4. K. Ceszei, N. Georganopoulos, Z. Turanyi, A. Valko, "Evaluation of the BRAIN Candidate Mobility Management Protocol", Proceeding IST Summit '01, September 2001.
5. G. Almes, et al, "A One-way Packet Loss Metric for IPPM", IETF RFC 2680, Sep. 1999.
6. C. Demichelis, P. Chimento, "IP Packet Delay Variation Metric for IPPM", IETF draft, <draft-ietf-ippm-ipdv-09.txt> (work in progress), April 2002.
7. G. Almes, et al, "A One-way Delay Metric for IPPM", IETF RFC 2679, September 1999.
8. ns-2 website, www.isi.edu/nsnam/
9. K. Fall and S. Floyd, "Simulation-based Comparisons of Tahoe, Reno and SACK TCP", Computer Communications Review, July 1996.
10. A. Delgado, A. Mihailovic, N. Georganopoulos, A. H. Aghvami, "Adaptation of Transport Protocols for an IP-Micromobility Scheme", Proc. ICC '01, June 2001.