Comparison between Semi-soft and Hard handover in a Cellular IP networks

Vedran Mikac, Alen Bažant and ŽeljkoIlić

Department of Telecommunications
Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
{vedran.mikac, alen.bazant, zeljko.ilic}@fer.hr

Abstract—**Due to the lack of the capacity per user connected to one access point (AP), the hotspot network providers are forced to constantly increase a number of APs to accommodate the customer's growing capacity needs. For that reason, APs cover a small area, so mobile hosts must initiate frequent handovers. During the handover mobile hosts can not receive data, which results in a packet loss and in performance degradation of higher layer protocols, especially TCP (Transmission Control Protocol). Cellular IP (CIP) is a protocol which supports semi-soft handover and shows a good performance of TCP in the throughput aspect. In this paper, as a part of our research we have implemented a CIP protocol in OPNET modeling environment [3]. We present simulation results showing handover performance and its influence on the higher layer protocols, especially TCP. Also, we have pointed out several reliability issues in the CIP and suggested improvements that correct mentioned problems. It is shown that CIP is an efficient solution in frequent handover environment.** *(Abstract)*

I   INTRODUCTION

Recent price cuts in wireless networking equipment resulted in emergence of numerous public area networks. Providers compete for customers through lower prices and/or better service. One of the aspects of better service are faster data transfer rates. Each access point (AP) has a very limited capacity, and to serve more than just a few users on a limited area, with a decent capacity, more than one AP is required. Large number of APs covering small areas result in an increased number of handovers for mobile users [9]. During the handover mobile host cannot receive data, which results in a packet loss. Since Transmission Control Protocol (TCP) is coded with the presumption that all data loss in the network is due to network congestion, it initiates congestion control algorithms that decrease data sending rate, and unnecessarily degrade communication between two correspondent stations. This degradation makes real time services, such as Voice over IP (VoIP), unusable.

This paper compares semi-soft handover and hard handover, in Cellular IP (CIP) networks [8], where TCP is used as a transport layer protocol. It presents results that favor semi-soft handover as a better alternative to the hard handover in a system that does not support soft handover. As this paper is a part of a work in progress at the end it focuses on proposed improvements which are to be implemented and tested.

The reminder of this paper is organized as follows. Second section is split into two subsections. The first one briefly explains architecture of a CIP network, while the second one explains handover procedure. The first part of the third section focuses on the implemented scenarios, while the second part compares semi-soft and hard handoff. Fourth section is dedicated to the CIP improvement suggestions and is followed by the conclusion.

II   CELLULAR IP (CIP)

A.   CIP Network Architecture

Cellular IP (CIP) is one of the micro-mobility protocols and typical usage for it is in a public area network. It can be used in WANs, but then it needs to interact with Mobile IP (MIP) for wider area mobility. Typical CIP network has a tree hierarchy, where the root node is CIP gateway. Leaves of the tree are APs, which are connected to the root node via one or several CIP switches. Nodes that are closer to the leaves of the tree are downlink neighbors to the neighbor nodes that are closer to the CIP gateway. Fig. 1 represents typical CIP network. Because CIP was intended to work in relatively benign environments, there are no redundant connections between CIP switches.
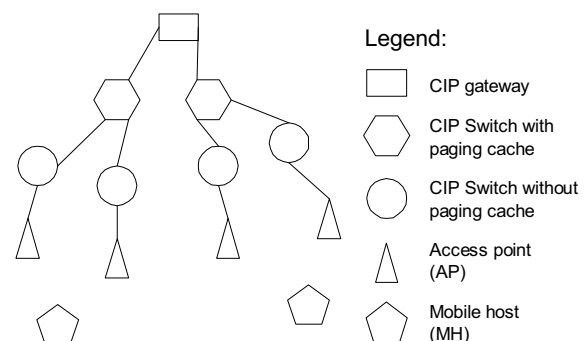

Figure 1. Typical CIP network

Every CIP switch must have a routing cache and can have a paging cache. Generally speaking, routing cache is

used to forward packets to the MH that is actively sending and receiving data. Paging cache is used to forward packets to the stations that have not been active for a specified time. This concept has been borrowed from cellular networks, where, the most of the time, exact position of the MH is not important. CIP network only has to know approximate position (paging area) of a MH. This way packets that keep the connection alive and report station's exact position are not sent very often, so amount of signaling messages is greatly decreased. This is a very good property of the protocol, because it saves the mobile device's short lived battery. CIP standard draft recommends that station is declared inactive after 10 s.

When forwarding packets from CIP gateway to MNs, CIP switches first look for entries in the routing cache. If no entry was found, than paging cache is being browsed. If no entry is found there packet is discarded, otherwise, if packet's destination is found in one of these two caches, it is forwarded to the specified downlink neighbor. If there is no paging cache and an entry for specified station is not found in the routing cache, packet is broadcasted to all downlink neighbors. Before sending any data MH must send route update packet to the gateway. Route update packet creates route cache entries in all CIP switches it passes through. That way, CIP switches can forward data down to the MH. Both paging cache and route cache entries have to be refreshed. To decrease the number of signaling packets, every packet that MH sends refreshes timers in paging and route caches. There are also paging update packets, which create entries in the paging caches on the way to the gateway node. These packets have to be sent when the mobile node is connecting to the CIP network, changes the paging area or before it's paging cache entry expires. These packets do not have to be sent very often, because paging area consists of several adjacent CIP switches and APs. Paging cache entries also expire after longer period of time, so there is no need to constantly refresh them.

When one MH is sending packets to another MH in the same network all packets must be routed through the gateway. Routing scheme where all the data must pass through the gateway node has several disadvantages. Obvious disadvantage is that data between two mobile nodes does not have to be sent through topmost crossover node (gateway), because there is a great probability that crossover node exists closer to the correspondents. Also, in a larger CIP network, links that connect gateway to the CIP switches must have an immense capacity. That poses a great financial problem to a CIP network owner that must rent capacity. As it was mentioned earlier, interaction between MIP and CIP, to ensure WAN mobility, takes place in the gateway node, which can have a foreign agent functionality (if MIPv4 is being used).

B. CIP handover procedure

There are two kinds of Layer-2 handover. First one is hard handover. This kind of handover is not recommended, because MH can not receive data during handover period. This kind of handover is associated with WLAN and GSM networks. Soft handover is associated with UMTS networks [7]. During this handover station receives data even during handover (station can receive data from more than one station at once). CIP standard draft introduces semi-soft handover. Idea behind semi-soft handover is minimization of packet loss, so as little as possible or even zero packets are lost during the handover. Semi-soft handover is in essence a hard handover, but it has lesser impact on the higher layer protocols (in our case TCP) as our results will show. Semi-soft handover announces that the MH will perform handover before actual hard handover. It sends route update packet, which is propagated to the first crossover node, where it creates second route cache entry for the same MH. The second route cache entry is used to send packets with slight delay on the new route. In our simulation we have suggested that the time delay should close to the handover time. This way, during the handover no packets are lost, and TCP does not initiate Karn's algorithm [2] which slows down data transfer. After the hard handover, MH sends another route update packet, which tears down old route, so all the data is being sent only on the new route. If the crossover node isn't the first uplink neighbor of both APs, there is no need to remove all route cache entries on the old routing path. Only entry that must be removed with route teardown package is the entry in the crossover node's routing cache. With the entry that forwards packets on the old route removed, other route cache entries on the old route will shortly expire.

III SIMULATION RESULTS

A. Implementation Scenario

In the draft [1], CIP switches are considered to have functionality of a switch and an AP. In our implementation we have special nodes that have switching functionality and special nodes that have AP functionality. This is more practical approach, because cheaper APs usually do not have more than one Ethernet port. For the simulation purposes, after MH is turned on, it associates with the closest AP. In reality, when MH is turned on it scans the area (active scanning), and associates with an AP that has the best signal-to-noise ratio (SNR). In our simulation we assumed that when MH gets a beacon frame from an AP, it measures frame's SNR and stores that information in an internal cache along with the sending AP MAC address. Every time the beacon frame is being received, MH browses through the cache, and determines which AP has best SNR. To avoid pig-pong effect, handover procedure is not initiated until an AP, other than current AP that MH is being associated with, has better SNR ratio for some time.

For the simulation purposes we have crated two scenarios shown in Fig. 2, which are identical in all means except that in first one we implemented hard handover, and in second one we performed measurements with semi-soft handover. To evaluate TCP performance, Workstation node sends a very large file to the MH using File Transfer Protocol (FTP). In our simulation sending

file has approximately 300MB. It takes about 218 second transmit file this big through 11 Mbit link (capacity of the wireless link, considered from FTP's layer is much lower). That way we have ensured that during the simulation which lasts 85s there will be constant data flow between Workstation and MH.
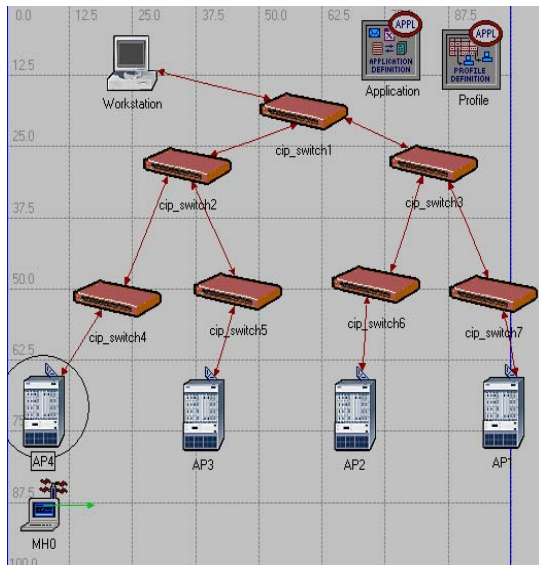

Figure 2. Topology of both simulation scenarios

B. Comparison and Conclusions

Fig.3 shows total amount of data received at the MH. It can be noticed that when using semi-soft handover more data from high layer is transferred.
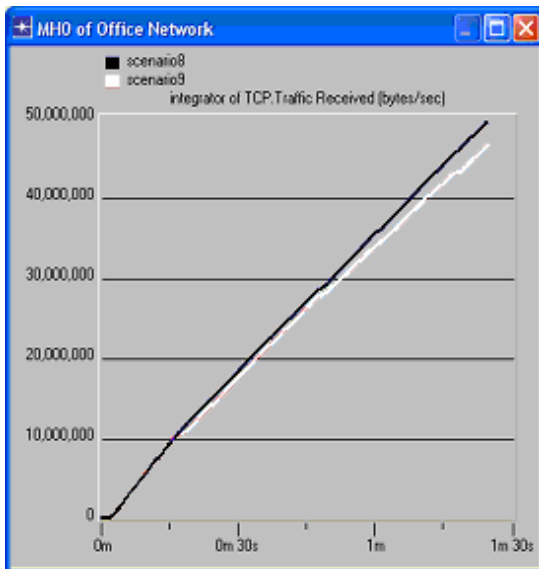

Figure 3. Total traffic received

At the end of simulation, total amount of data transferred when using semi-soft handover is 49.46 MB (dark line). In a scenario that uses hard handover the total

amount of data transferred is 46.52 MB (light line). Difference between this two scenarios is caused by TCP congestion avoidance algorithms. Fig. 4 shows AP throughput when hard handover is being used, and Fig. 5 shows AP throughput when using semi-soft handover.

Figure 4. AP throughput with hard handover

It is obvious that during the hard handover, there are breaks in the data transfer. When using semi-soft handover, data is being constantly sent. Number of lost packets is minimal, so TCP at the sender's side does not decrease congestion window. As it was mentioned before, during the handover crossover CIP switch sends data on the new route with a delay.
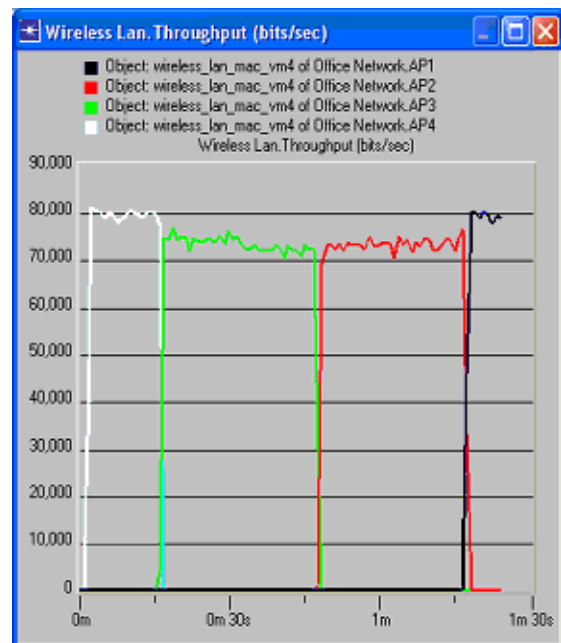

Figure 5. AP throughput with semi-soft handover

For this paper we haven't studied effects of the delay duration. If VoIP would be used, it is clear that this handover delay should be as small as possible, so parties involved in a conversation would not experience increased delay effects. On the other hand, if the delay is too short, increased number of packets will be lost during the handover, so applications using TCP as the transport layer protocol will have a slow connection. For this reason we have proposed that the delay should be little longer than the average handover time. Every switch or router that packet passes through adds to total delay. With this taken into consideration, maximum semi-soft handover delay should be much less than 200ms limit.

IV   PROTOCOL IMPROVEMENT SUGGESTIONS

Authors of CIP draft did not consider protocol reliability as one of the crucial issues. As it was mentioned before, CIP is intended for environments where the probability of a link failure is negligible. Problem appears in the wireless part of CIP network. We suggest that acknowledgment (ACK) messages should be introduced. Currently if route or paging update package is lost, there is no retransmission of the package. We believe that ACK messages are crucial, because IEEE 802.11 operates in a Industrial-Scientific-Medical (ISM) frequency band. Other devices that also use ISM band could interfere with CIP signaling packages [5].

Also, we suggested that there should be some kind of package numbering implemented. Each package would get a sequence number when passing through gateway. This sequence number would be written in an option field of IP header. Furthermore, BS should buffer several sent packages. This way, if a package was lost due to interference with other devices on the same frequency band, it could be quickly retransmitted, before TCP retransmission takes place. Downfall to this idea is that all TCP packets already have a segment number, and that CIP sequence numbers would be unnecessary data redundancy. On the other hand, many real time applications use UDP, where there is no segment numbering [6]. Basic problem with this idea is that it would make CIP more complex and slower.

Every CIP node knows it's uplink neighbor. If a link to the uplink neighbor fails, there is no redundant link data could be transferred through. It would be more than an inconvenience for a provider if a link that connects whole paging area to the gateway would fail. We suggest that every link that connects a node to it's uplink neighbor should be tested with as short packets as possible. If an uplink node does not respond with an ACK in a short period of time, this link would be considered down. To remedy that situation, we suggest that all the nodes that have same number of hops to the gateway node should be interconnected. When the link from the first node to the uplink neighbor fails, all the traffic directed to the gateway should be sent using second node's uplink neighbor, as shown in Fig 6.
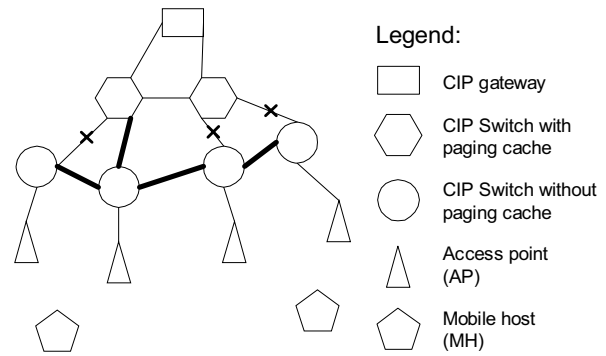

Figure 6. Suggested CIP network topology

We believe that this additional signaling would have a negligible impact on CIP network because one Gigabit Ethernet link can serve large number of APs.

V   CONCLUSION

Our results show semi-soft handover is superior to the conventional hard handover. It should be noted that sending same packets twice does take a lot of capacity, since handover period is relatively short, and we believe that this is not a major concern. Protocol CIP is a good addition to the Mobile IP for micro-mobility because it offers fast handover with almost no packet loss. Main problem with CIP is that is has no lost packet recovery mechanism. Multipath fading, interference with other devices, shadowing and other phenomena are a reality in a wireless medium. Protocol CIP should be enhanced with mechanism that can provide more reliable data transfers. TCP does offer reliability, but it has a great cost in a transfer rate when several consecutive packets are being lost. We plan to implement our suggestions, and hope to improve CIP so it could be used in real systems.

REFERENCES

[1]   Campbell, J. Gomez, C-Y. Wan, Z. Turanyi, A. Valko. Cellular IP. *Internet Draft*, IETF, 1999

[2]   W. E. Stevens, 1995. *TCP/IP Illustrated: The Protocols, Volume 1.* Adisson Wesley Longman Inc.

[3]   http://www.opnet.com

[4]   Opnet Modeler Documentation

[5]   Matthew S. Gast. *802.11 Wireless Networks: The Definitive Guide.* O'Reilly & Associates, Sebastopol, CA, April 2002.

[6]   RFC 768. User Datagram Protocol. August 1980.

[7]   Harri Holma and Antti Toskala. WCDMA for UMTS – Radio Access for Third Generation Mobile Communications. John Wiley & Sons, Ltd. 3rd edition. 2004.

[8]   A. T. Campbell *et al.*, "Design, Implementation, and Evaluation of Cellular IP," *IEEE Pers. Commun.*, vol. 7, no. 4, Aug. 2000, pp. 42–49.

[9]   K. Pahlavan *et al.*, "Handoff in Hybrid Mobile Data Networks," *IEEE Pers. Commun.*, vol. 7, no. 4, April 2000, pp. 34–47.