MULTIMEDIA FLOW MOBILITY IN HETEROGENEOUS NETWORKS USING MULTIHOMED MOBILE IPv6

Christer Åhlund, Robert Brännström,
Karl Andersson and Örjan Tjernström [1])

Abstract
Communication in next generation networks will use multiple access technologies, creating a heterogeneous network environment. To enable end-user terminals to move between access networks with minimal disruption, the terminals should be able to maintain multiple active network connections. Such a multihomed mobile host will experience different capabilities and coverage area depending on the access technologies. This paper proposes and evaluates an extension to Mobile IPv6 enabling multihoming, regardless of the access technology. Mobility of multimedia communication in this environment should adapt to changing conditions and be base on dynamic measurements and user preferences. The proposed architecture gives opportunities for mobile multimedia applications to use multiple access networks simultaneously and the possibility to move individual flows. Media flows are identified by destination IP address, protocol and port number. Results from a real world prototype are presented using three different wireless access technologies: 802.11, UMTS and 802.16-2004.

1. Introduction

In the new generations of wireless networks, seamless mobility across heterogeneous networks will be supported. A widespread vision of the fourth generation (4G) mobile networks or Next Generation Networks (NGN) includes coexistence of current wireless technologies such as WLAN, WiMAX, General Packet Radio Service (GPRS) and Universal Mobile Telecommunications System (UMTS). Different technologies will be bound together into a single network and the IP will be the glue. The mobile nodes (MN) will be equipped with multiple access network cards and users will be able to roam transparently over networks in a seamless manner. Software defined radio have however recently gained new interest in research communities and may in the future be an alternative to using multiple interfaces.

To manage mobility between access points (APs) there are basically two approaches, namely mobile-controlled handover (MCHO) and network-controlled handover (NCHO). NCHO requires network providers to manage handovers between different technologies and is not a feasible solution in today’s heterogeneous networks. On the contrary MCHO can easily be adopted by adding multiple interfaces to MNs and software upgrades by users of this functionality. For heterogeneous networks, policies are required for the decision of which type of technology to use in different situations. These policies should be based on user preferences as well as continuously evaluating performance of available access networks.

[1] Luleå University of Technology, LTU Skellefteå, SE-931 87 Skellefteå, Sweden

The ideas and terms of an IETF proposed policy model [14] are widely spread. Figure 1 illustrates a policy model that is based on the IETF proposed Policy Decision Point (PDP) and Policy Enforcement Point (PEP) entities, extended with a Policy Repository (PR) entity.

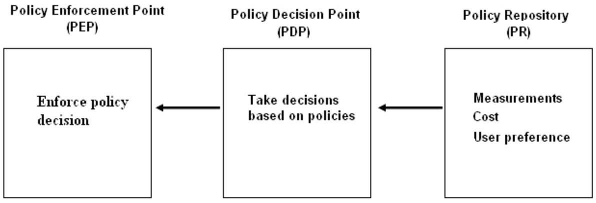

Figure 1. Policy-Based Decision Model.

As figure 1 show, the model consists of three entities:

- A Policy Repository: The PR is responsible for delivery of requested policy parameters to the Policy Decision Point. The PR contains information such as user preferences, signal strength and cost of available access networks. The PR can obtain information through measurements of the environment.
- A Policy Decision Point: The PDP is the control entity that evaluates access networks through policy decisions. The policy decisions are based on the parameters received from the PR. If the PDP decides that a handover is motivated, the PDP informs the PEP to perform a handover.
- A Policy Enforcement Point: The PEP receives policy decisions from the PDP and performs the actual handover. The PEP is said to enforce the policy decision [8].

The physical location of the PDP separates policy systems to perform an MCHO or an NCHO.

To manage mobility for an MN connecting to IP networks, where applications and users are unaware of the network mobility, Mobile IP (MIP) is deployed. The MIP architecture incorporates a home agent (HA), and the MN. An MN connected to the home network will operate according to normal IP network operations, without using MIP. When an MN connects to a foreign network it will register its new location by sending a binding update (BU) containing the current address to the HA. This address is called the care-of address (CoA). The registration sent by an MN to the HA will create a binding in the HA between the home address (HoA) and the CoA.

When packets to the MN are discovered at the home network, the HA will forward the packets to the CoA using tunneling. A tunnel encapsulates the received packet as a payload in a new packet with an outer IP header having the CoA as the destination and the HA as the source. When the packet arrives at the MN, it will be decapsulated by the networking software. The outer packet header is removed before the packet is handed to upper layers. In MIPv6 a packet sent to a Correspondent Node (CN) will use the MN’s CoA as the source, and the HoA will be added in the home address destination option. Since the addresses are topology-correct, ingress filtering is avoided. The CN receiving the packet will replace the source address with the address in the home address destination option before handing the packet to the transport layer. The routing created by MIP is referred to as triangular routing (see figure 2). Here packets from a CN are sent to the HA.

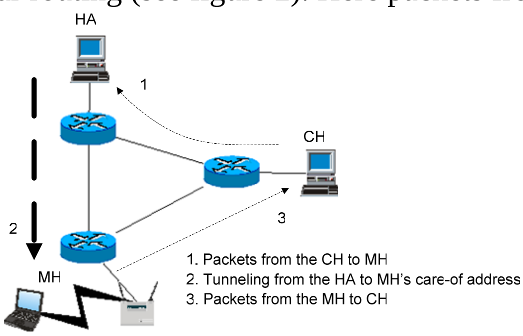

Figure 2. Triangular routing in Mobile IP.

The HA tunnels packets to the MN, and the MN sends its traffic directly from its current location to the CN, making a triangle. In the case of CN being unaware of MIP the traffic must be tunneled from the MN to the HA and then forwarded by the HA to the CN.

To optimize routing between the MN and a CN, route optimization is used (see figure 3). An MN receiving packets via the HA informs the CN about its current CoA in a BU. When a CN receives a BU it will start to send packets directly to the MN using the CoA as the destination address.

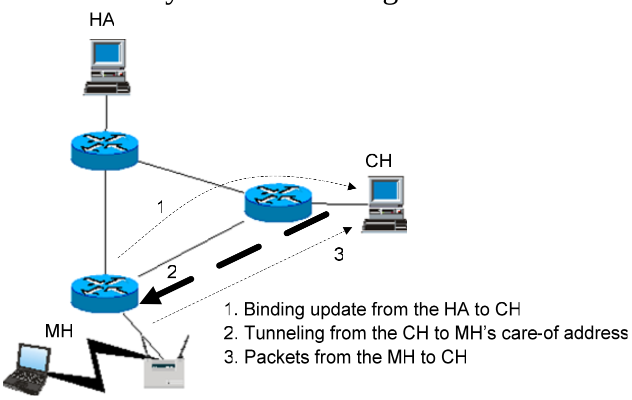

Figure 3. Route optimization in Mobile IP.

In MIPv6, support for route optimization is built into IPv6. The CN uses the IPv6 routing header, where the destination of the packet is the MN's CoA and the address in the routing header is the MN's HoA. When the MN receives such a packet, the destination field will be updated with the MN's HoA before handing the packet to the transport layer.

To secure MIPv6 route optimization the return routability procedure is used (see figure 4). In the procedure four messages are sent; Home Test Init (HoTI), Care-of Test Init (CoTI), Home Test (HoT) and Care-of Test (CoT).

Before the MN sends a BU to the CN it sends a HoTI message through the HA. A CoTI message is also sent directly to the CN according to IP routing. When receiving these messages the CN responds with the HoT and CoT messages, where HoT is sent through the MN's HA and the CoT message is sent directly to the MN's CoA. The MN derives a binding management key from the information in the HoT and CoT messages. After this, the BU will be sent to the CN. The CN will derive the binding management key from the information in the BU.

The return routability procedure verifies that the MN is reachable both through its HoA and its CoA.

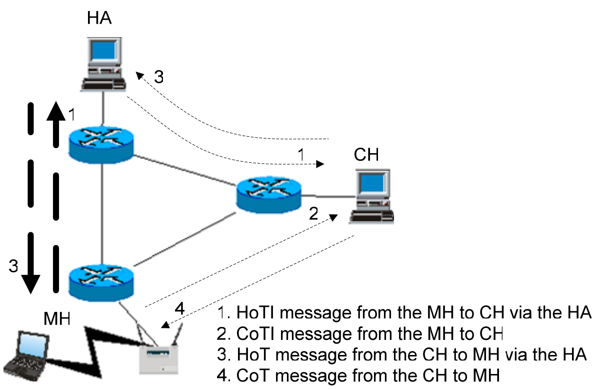

Figure 4. Return routability in Mobile IP.

To secure the information exchanged in the return routability procedure, IPSec [6] can be used between the MN and its HA for the HoTI and HoT messages. A malicious node has to intercept both HoT and CoT messages to create the binding management key. Return routability is required each time the MN changes its CoA. As long as the same CoA is used the same binding management key is valid.

The MIP solution is attractive since it enables mobility with the most widely used protocols at the transport layer, the Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP). With MIP, protocols above the network layer are unaware of network mobility. However, one problem with MIP is the registration time when moving between networks, especially if there is a long distance between the HA and the foreign networks visited by an MN. MIP will probably be most used with MNs connecting wirelessly and this may cause problems because of rapidly changing conditions in the wireless network. An MN switching between APs connected to different networks will require a new registration each time. The time it takes performing a handover may cause UDP packets to be dropped and TCP flows to break.

The rest of the paper is organized in the following way. Section 2 describes the port-based MIP architecture. Section 3 presents conducted experiments. Section 4 describes related works and section 5 presents conclusions and future work.

2. The Port-based Mobile IP Architecture

In the scenario shown in figure 5, an MN should be able to send traffic flows via different interfaces. In the proposed standard for MIPv6 a MN disconnecting from its home network (using one HoA) can only use one wireless connection (i.e. register one CoA) at a time. In the multihomed extension for MIP, M-MIP [15] multiple CoAs are managed. With M-MIP, CNs can associate the MN's HoA with multiple CoAs. In the case of with two registered CoAs the HA and CN may use different CoAs to reach the MN. However, M-MIP do not enable a CN or the HA to control the use of multiple CoAs by itself. Using multiple CoAs is beneficial if the total amount of traffic capacity needed extends the capability of one single interface. In that case flows can be sent via different interfaces.

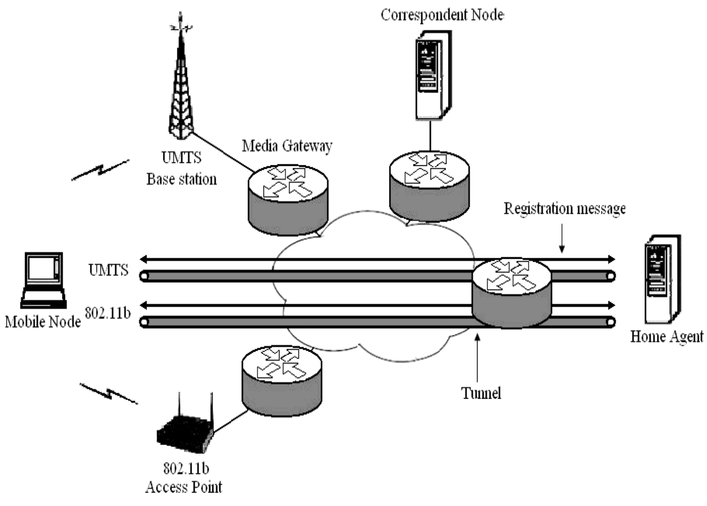

Figure 5. System overview when the MN is connected to both an 802.11b and a UMTS network.

To enable port-based MIP we extend the M-MIP proposal to include a flow mobility option header, specifying the protocol and port number when registering a binding. By doing this an MN can register a binding that informs the CN or the HA that only a single flow shall be forwarded to the specific CoA. To control multiple flows the MN can include several flow mobility option headers in the BU. Beyond enabling flow mobility at the network layer the extension also enables flow mobility between devices. If a user register multiple devices with the same HoA (e.g. a phone and a laptop), it is possible to redirect flows between the devices.

The modifications consist of two flags added in the BU message and a new option header hosting the protocol number and the port number. Figure 6 illustrates the modified BU header.

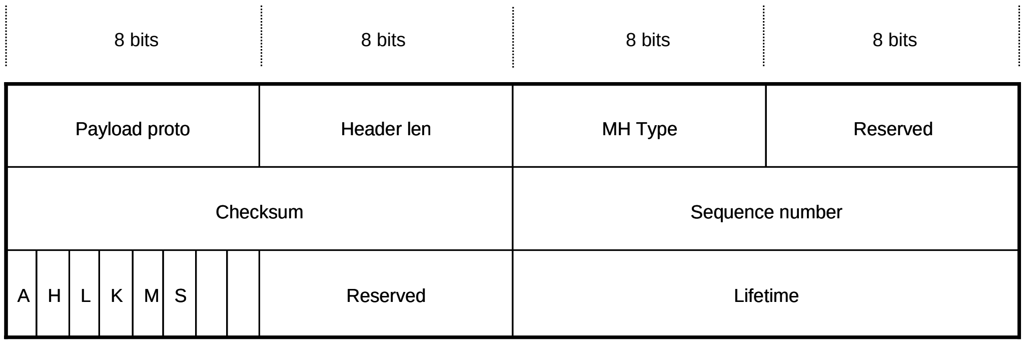

Figure 6. Binding update header with M and S flag.

The flag named the M-flag indicates a multihomed binding. This means that with the M-flag, currently registered bindings will be kept and without the M-flag they will be deleted. The S-flag is used by the MN to inform the HA and CNs of which CoA to use as default. Figure 7 illustrates the flow mobility option header.

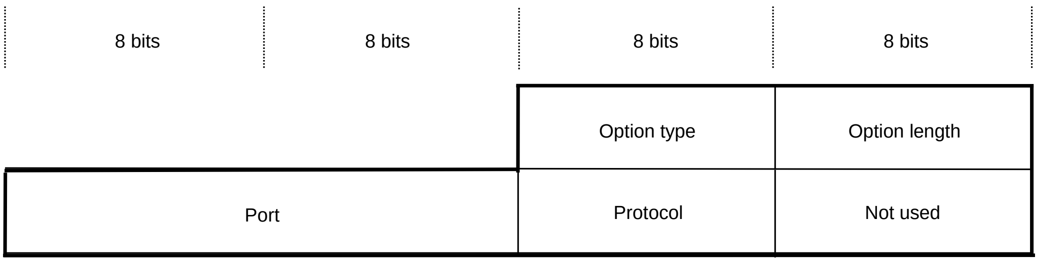

Figure 7. Flow mobility option header.

The port field identifies the destination port at the MN. The protocol field represents the transport protocol number. In figure 8, bindings are shown for both an IP address as well as for a protocol and a port.

HoA	**CoA**	**Protocol**	**Port**	**Lifetime**	**Flags**
3ffe::a:b:c:d	3ffc::1:5:a:b:c:d	-1	-1	150	A/H/L/K/M/S
3ffe::a:b:c:d	3ffc::1:6:a:b:c:a	6	6935	200	A/H/L/K/M
3ffe::a:b:c:d	3ffc::1:a:a:b:c:d	17	7830	150	A/H/L/K/M

Figure 8. Binding cache.

The proposed solution requires the network layer to look for port numbers in the transport header. This is however nothing unique and is for example used to enable fast forwarding and to filter packets in access control lists. With wireless communication and mobility, information between layers is needed to create solutions that scale. A new research area working with these questions is cross-layer communication.

When an MN discovers a foreign network, it acquires an IP address and registers the CoA with it's HA. If this is the first registration a BU is sent without the M or S flag. A BU without the M-flag means that previously added registrations are deleted and that this binding is the one selected (without using the s-flag).

If a second foreign network is discovered, another registration is sent to the HA. In this registration the MN adds the M-flag in the BU. When registering a new interface it should not be selected until an evaluation of the interface is conducted, assuming a previously added interface is operational. If the new interface performs better than the interface previously used, a new binding update is sent for the new CoA with the M- and S-flag.

Without the option header adding protocol and port to the BU, all traffic (from the same or other CNs) is sent to the same CoA. By adding such an option, a single flow can be redirected to another interface (CoA). If e.g. the WiMAX interface is used and traffic from CNs via the HA congests the WiMAX connection, one or more flows should then be moved to an alternative interface, e.g. WLAN. Such a scenario could be an e-meeting [7], where voice communication requiring rather strict jitter and delay values should be given high priority and kept at the WiMAX interface, while video showing presence of participants can be given lower priority and be moved to an WLAN interface. In this case a BU is sent to the HA with the option header informing the HA of what protocol number and destination port number to be redirected to another CoA.

For route optimization a BU is sent to the CN. This BU can be valid for all traffic sent from the CN or just for a specific flow. This means that some traffic may go via the HA and some traffic can be sent directly.

In the case of all traffic being redirected, a BU is sent to the CN without the S- and M-flags and without the flow mobility option. In this case all traffic is sent from the CN to the registered CoA. To direct a flow from a CN to another CoA a BU is needed using the M- and S-flags as well as the flow mobility option.

For each new CoA, return routability is invoked. No changes in messages are needed except for the added option and the two extra flags in the BU. Return routability is only needed when adding a new CoA. In the case of handover for specific flows (by adding the flow mobility option to the BU) to an already registered CoA, no new return routability needs to be invoked.

3. The prototype

The prototype has mainly been tested on a laptop running the Linux distribution Fedora core 4, with kernel 2.6.11, Java™ 2 Runtime Environment, Standard Edition v 1.5.0 (J2SE). The Linux kernel was compiled to support the following network devices: Universal TUN/TAP device driver, PPP (point-to-point protocol), IP advanced router and IP policy routing. The open source packages openVPN and iproute2 were installed. The prototype uses iptables to mark packets for policy routing. This enables selection of packets to an MN with a specified protocol and port. An ip rule directs the packets to a dedicated routing table.

In this prototype all traffic is sent via the HA. For each added CoA, a tunnel is established between the HA and the MN. The selection of what interfaces to use is based on the policy value expressed in formulas (1-4) shown below. A detailed explanation of the formulas can be found in [2].

To estimate a network's capacity, the Relative Network Load (RNL) is calculated in the MN. RNL represents a quality value for each network based on round trip time (RTT) and jitter values. $\bar{z}_n$ is the mean value of RTT metrics (rtt_n) for MIP registration messages between the MN and its HA. $\bar{x}_n$ is the mean value of times between arrivals of MIP registration messages at the MN, and V_n is the variance between these messages. The variable h determines the size of the history window for the weighted average calculations. For example, when h=5 the most recent value will contribute 20 per cent to the calculated $\bar{x}_n$, $\bar{z}_n$ and V_n values.

$$\bar{z}_n = \frac{1}{h} rtt_n + \frac{h-1}{h} \bar{z}_{n-1} \qquad (1)$$

$$\bar{x}_n = \frac{1}{h} \delta_n + \frac{h-1}{h} \bar{x}_{n-1} \qquad (2)$$

$$V_n = \frac{1}{h} (\delta_n - \bar{x}_n)^2 + \frac{h-1}{h} * V_{n-1} \qquad (3)$$

$$RNL_n = \bar{z}_n + V_n \qquad (4)$$

The variables h, $\bar{x}_0$, and V_0 are initialized with the following values:

$\frac{1}{h} = \{z : z > 0 \wedge z \leq 1\}$

$V_0 = 0$

$\bar{x}_0$ = defined advertisement time.

The variable δ_n is calculated as: $\delta_n = \{t_n - t_{n-1} : n > 0\}$

Where $t_n - t_{n-1}$ is the time difference between consecutive MIP registration messages received at the MN.

To evaluate the prototype, three access network technologies were used; WLAN (802.11b), WiMAX (802.16-2004) and UMTS. The topology is shown in figure 9. To simulate a Skype video call of 20 kbps (UDP) we used the Iperf traffic generator.

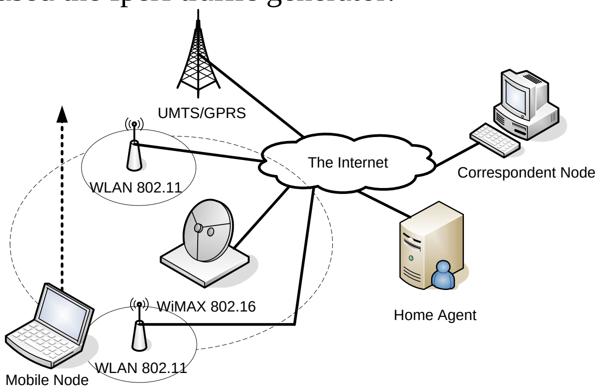

Figure 9. Evaluation topology.

Traffic from the CN is sent to the MN's home network, where the HA intercepts packets destined for the MN and tunnels them to the selected CoA. For each registered CoA a tunnel is installed. Experiments conducted used h = 5 (in formulas 1-3). The movement pattern during the experiment is as follows: the MN starts at a place nearby an 802.11 AP and moves towards another 802.11 AP. Between these two APs there is bad WLAN coverage. Beyond the second 802.11 AP, both 802.11 and 802.16 have bad coverage. The 802.16 cell cover the area of both 802.11 cells and the UMTS network cover both the 802.11 and 802.16 cells. After some time using the UMTS access the MN turns and starts to approach the (second) 802.11 AP again. Results are presented in figure 10.

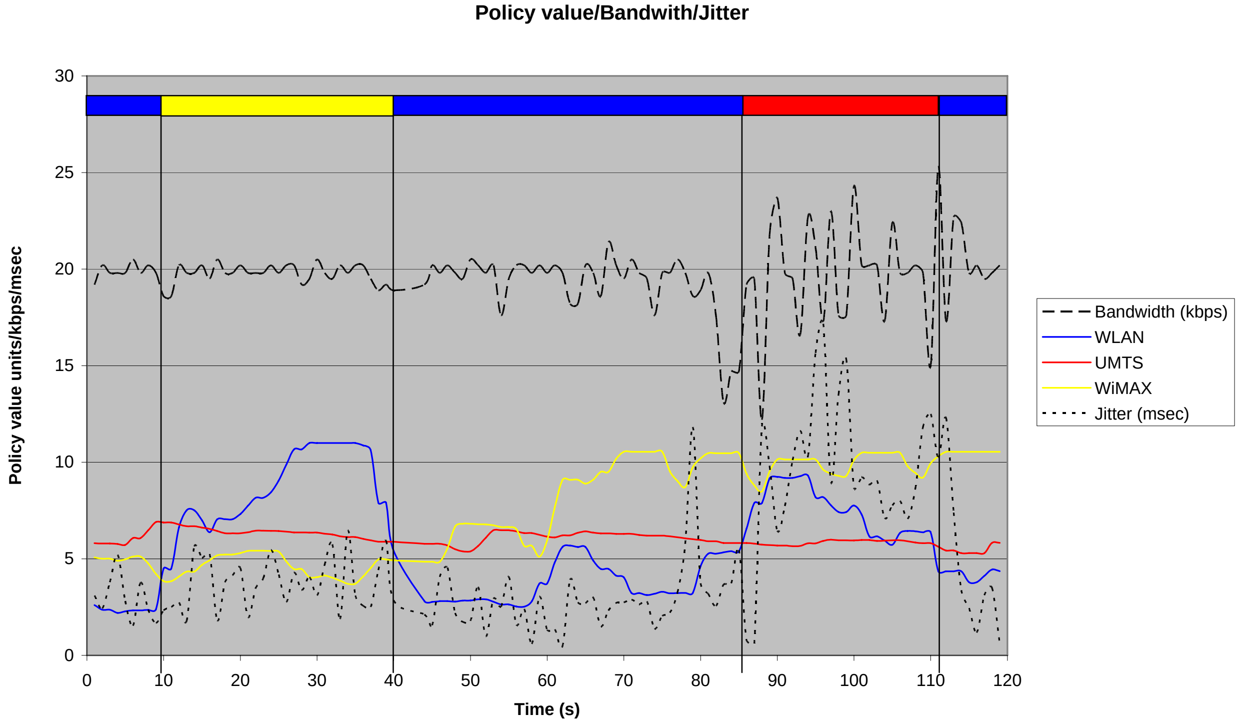

Figure 10. The calculated policy values for each access network at the MN. The bandwidth and jitter is also shown.

Each vertical line illustrates the time when a handover was performed. The blue curve illustrates the policy value for the 802.11 access, the red curve illustrates policy value for the UMTS access and the yellow curve plots the policy value for the 802.16 access. The access network having the lowest policy value is selected. The upper black plot shows the bandwidth (kbps) received at the MN and the lower black plot shows the jitter (msec) of the received traffic. The CN send constant bit-rate traffic (UDP) of 20 kbps. The 802.11 APs have a capacity of 11Mbps (5.5 Mbps in practice) and the 802.16 AP was set to 2.22 Mbps (BPSK ½ modulation). The UMTS network performed about 300 kbps downlink (depending on other ongoing traffic).
The evaluation started indoors and the first 802.11 AP was selected the first 9 seconds. At that time the MN's distance to the 802.11 AP render in an increasing policy metric due to increased RTT and jitter metrics of MIP registration messages. At this time the 802.16 policy metric is the lowest and therefore is selected. The handover renders in a small increase in jitter metrics but the bandwidth is kept stable with small fluctuations. The short dip in bandwidth at 10 seconds from start indicates that the handover decision is taken at the right time.
After 40 seconds from start the connection to the second 802.11 AP is considered better than the 802.16 access regarding RTT and jitter metrics. This is illustrated by the jitter plot showing lower jitter metrics for the received traffic, after handover from the 802.16 access network to the 802.11 access network. The bandwidth is shown to be stable during this handover as well. However, a small increase in fluctuation of measured bandwidth is experienced.
After 83 seconds from start, the MN leaves the indoor environment (where the 802.11 and 802.16 access networks are installed) and enters an outdoor environment. Handover to the UMTS access network takes place after 86 seconds from start due to both the 802.11 and 802.16 radio signals being faded by the wall of the left building. Since today's UMTS networks prioritize circuit switched traffic before packet switched (IP) traffic, the jitter increases as well as the fluctuations of both the bandwidth and the jitter metrics. The trend of the bandwidth and jitter plots shows that the handover decision is taken at the right moment, since this trend is kept while using the UMTS

network. When returning back into the building (after 111 seconds from start) handover takes place to the 802.11 access network, rapidly increasing both the bandwidth and jitter metrics.

4. Related work

Methods for horizontal and vertical handovers are discussed in [9,12]. These approaches use multicast to reach multiple nearby APs. MNs instruct APs to forward or buffer data packets for it. If not delivered to the MN, these packets are dropped after some time. Our solution uses multihoming to maintain multiple registrations and by this avoids forwarding between APs.

Soliman et al [10] present a proposal to lower the delay with MIP messages and thereby manage handover at the network layer more efficiently, considering the time for handovers. The proposal uses two care-of addresses; link local care-of address and regional care-of address. In our solution the possibility of maintaining multiple bindings enables a MN to perform soft handovers.

A solution for fast handovers is presented by Dommety et al [4]. It uses signalling between the MN, the old AP and the new AP entered to avoid losing packets. Packets will be forwarded from the old AP to the new AP in order to avoid packet losses. Our solution avoids forwarding between APs and complicated signalling between access networks possibly owned by different providers

Hseih et al [5] combine the proposals [10] [4] and extend it to reduce the handover time even further. The handover time in this work is the same as the handover times for datalink layer handovers.

Wang et al [13] present a user-based policy for determing the currently best available access network in a heterogeneous network. The bandwidth is monitored and announced by APs so the MNs can calculate the utilization of each AP. Other parameters such as capacity and cost also affect the policy decisions. In our solution a new parameter (RNL) is used in the policy model for selection of which access network to use.

Chen et al [3] propose a Smart Decision Model to determine the best available network. The decision model considers factors such as user preferences, system information and properties of available access networks. The models score function is described in detail.

Bi et al [1] propose an integrated IP-layer handover solution that targets the IP-layer handoff delay. Policies use criterion from user profiles, service requirements and network environment. An adaptive handover control scheme combines probed and monitored (dynamic and static) information. Cross-layer signalling (e.g. L2 triggers) enhances the IP-layer handover.

Soliman et al [11] propose two flow movement options in MIPv6. One is based on IP addresses, protocol and ports and the other is based on the IPv6 flow label. The ideas are similar to ours but the authors do not describe how to manage multiple simultaneous bindings. Our proposal is based on multiple bindings (multihoming) and adding an option based on the flow label is only depending on the implementation possibilities to mark packets based on the flow label.

5. Conclusion

In this paper, we described how Mobile IPv6 can be extended to handle port-based multihoming. By such extension different flows can be destined through different interfaces on the MN leveraging differences in coverage, Quality of Service, cost, bandwidth, delay, et cetera among different wireless and fixed access networks. Our results presented in this paper show that load and capacity can be measured on the different interfaces. The results from the measurements can be used to perform load balancing between different interfaces, both between the MN and multiple CNs and multiple flows between MN and a single CN. In a near future, we intend to investigate load balancing between multiple flows and a single CN further.

We intend to develop our ideas even further by exploring and introducing different sorts of cross-layer signalling, extending and integrating the policy-based decision model, and implementing pilots where new innovative services like mobile IP telephony and pervasive games are evaluated. Hand-over between different mobile access network technologies will be studied in detail. Particularly, components from the 3GPP IMS (IP Multimedia Subsystem) will be considered.

Acknowledgments

The work presented in this article is based on results from the HybriNet@Skellefteå project supported by Skellefteå Kraft. The prototype is developed in collaboration with TietoEnator, Telecom & Media in Skellefteå.

6. References

[1] Bi, Y., et al., ”An integrated IP-layer handover solution for next generation IP-based wireless network”, Proceedings of the 60th IEEE Vehicular Technology Conference, 2004.

[2] Brännström, R., “Network-layer mobility in wireless ad hoc access networks”, Licentiate in engineering thesis, Luleå University of Technology, December 2005. ISSN 1402-1757 / ISRN LTU-LIC-05/68-SE / NR 2005:68.

[3] Chen, L.-J., et al., “A smart decision model for vertical handoff”, In Proceedings 4th ANWIRE International Workshop on Wireless Internet and Reconfigurability, Athens, Greece, 2004.

[4] Dommety, G., et al., Fast Handover for Mobile IPv6, Internet Draft, 2004.

[5] Hseih, R., Zhou, Z.-G., and Seneviratne, A., "S-MIP: A Seamless Handoff Architecture for Mobile IP," IEEE INFOCOM 2003 - The Conference on Computer Communications, pp. 1774-1784, 2003.

[6] Kent, S., Atkinson, R., “Security Architecture for the Internet”, RFC 2401, 1998

[7] Marratech, http://www.marratech.com/

[8] Murray, K., Mathur, R., and Pesch, D., “Intelligent access and mobility management in heterogeneous wireless networks using policy”, In ACM 1st International Workshop on Information and Communication technologies, pages 181.186, 2003.

[9] Seshan, S., Balakrishnan, H., and Katz, R., “Handoffs in Cellular Wireless Networks: The Daedalus implementation and Experience”, Wireless Personal Computing, vol. 4, pp. 141-162, 1997.

[10] Soliman, H. C., et al., Hierarchical MIPv6 mobility management, Internet Draft, 2004.

[11] Soliman , H., Malki, K.-E., and Castelluccia, C., Flow movement in Mobile IPv6, Internet Draft, 2003.

[12] Stemm, M., Katz, R. H., Vertical Handoffs in Wireless Overlay Networks Mobile Networks and Applications, vol. 3, no. 4, pp. 335-350, 1998.

[13] Wang, H. J., Katz, R. H., and Giese, J., “Policy-enabled Handoffs Across Heterogeneous Wireless Networks”, Proceedings of the Second IEEEWorkshop on Mobile Computer Systems and Applications, pp. 51-60, 1999.

[14] Yavatkar, R., Pendarakis, D., and Guerin, R., “A Framework for Policy-based Admission Control”, RFC 2753, 2000.

[15] Åhlund, C., Brännström, R., and Zaslavsky, A., “M-MIP: Extended Mobile IP to Maintain Multiple Connections to Overlapping Wireless Access Networks”, International Conference on Networking, 2005.