The Operation Mode Selection in FMIPv6

Bin LIU , Philippe MARTINS
GET/ENST/INFRES Department
ENST - Paris - Ecole Nationale Supérieure des Télécommunications
LTCI-UMR 5141 CNRS, 46 rue Barrault, 75634 Paris Cedex 13, FRANCE
Email: bliu@enst.fr, martins@enst.fr

Philippe Bertin
France Telecom Research and Development
4 rue du Clos Courtel – BP91226
35512 Cesson Sévigné Cedex, FRANCE
Email: philippe.bertin@orange-ftgroup.com

Abstract

The Fast Handover for Mobile IPv6 (FMIPv6) [2] aims at reducing the long handover latency in Mobile IPv6 [1] by fast movement detection and fast binding update. Furthermore, it also reduces packet loss by buffering before the real link layer handover takes place. Due to uncertain link layer trigger time, however, the buffering mechanism sometimes introduces unacceptable handover latency for real-time traffics such as voice over IP (VoIP). In addition, if it's the mobile node that makes the handover decision, FMIPv6 also suffers from uncertainty of handover target, which is unfavorable to TCP traffics such as FTP. In order to eliminate these negative effects, we propose that FMIPv6 should operate in predictive mode under the control of network if possible for TCP traffics. As to the VoIP traffic, in terms of the handover latency and packet loss requirements, a proposed hybrid mode is chosen adaptively. Numerical simulation results for the VoIP traffic demonstrate obvious performance improvement in terms of call drop rate and a new performance metric called packet cost.

Keywords- FMIPv6; Operation mode; VoIP; TCP; Handover

1. Introduction

In wireless IP networks, Fast Handover for Mobile IPv6 (FMIPv6) [2] has been accepted as a promising IP layer handover solution in IETF to solve the problems of long handover latency and high packet loss in Mobile IPv6 (MIPv6) [1]. It achieves much shorter handover latency and lower packet loss by informing the MN of the new AR's advertised prefix and validating the prospective new CoA from any duplication on the new link prior to the MN's handover. For this purpose, some additional messages are introduced such as messages Router Solicitation for Proxy (RtSolPr), Proxy Router Advertisement (PrRtAdv), Fast Binding Update (FBU). These signaling exchanges performed before actual link layer handover are aided by the implementation of link layer triggers, such as Link Going Down. However, whether the introduction of various signaling and link layer triggers in FMIPv6 is adapted for delay-constrained real-time traffic or throughput-sensitive traffic is still an open issue. In this paper, we follow a timing diagram methodology to identify conditions where the best performance can be achieved. In [3], Seung-Hee Hwang has gotten a similar result through mathematical analysis. We further this work by studying the influences of different operation modes in FMIPv6 to real-time traffics such as VoIP, and to throughput-sensitive traffics such as FTP. Based on these analytical results, for real-time traffics, we propose a new hybrid operation mode to realize an adaptive operation mode selection in terms of handover latency requirement. The following numerical simulations show that our proposal achieves lower call drop rate and lower packet cost for the real-time traffic than conventional FMIPv6 does when the network transmission latency is under a threshold. While for throughput-sensitive traffics, we propose that the FMIPv6 should operate in predictive mode under the control of network if possible.

This paper is organized as follows: the FMIPv6 protocol and handover timing analysis of the UDP traffic are described in section 2. In section 3, we study the handover timing of the TCP traffic when FMIPv6 is used as handover protocol. The handover performances of these two kinds of traffics are studied by numerical simulations in section 4. Based on the analysis of these results, we propose a hybrid operation mode for real-time traffics in section 5. The numerical simulation results show its improvements on downlink handover delay and packet cost. Section 6 is the conclusion of this article.

2. The timing analysis for the UDP traffic

In case of the scenario of an IEEE 802.11 WIFI network, the MN firstly performs a scan at any time while it connects to the current link to discover available APs. The scanning period which although gets a list of APs together with physical layer information such as signal strength, often brings forth unwanted connection disruption with current link. After scanning, the MN selects one or more APs and sends their AP-IDs to PAR in message RtSolPr. The PAR sends back the corresponding NAR's prefix, IP address and L2 address with message PrRtAdv. Then the MN gets enough information to formulate a prospective new CoA (NCoA) while it is still connected to the PAR's link. When the signal strength or signal quality is under the predefined threshold, a Link Going Down (LGD) trigger is created. This trigger makes the MN send FBU to PAR to authorize it to bind PCoA (or oCoA) to NCoA. On receipt of the FBU, the PAR establishes a tunnel with NAR and buffers arriving packets destined to PCoA. Before PAR sending an FBack to the MN, it sends the NCoA to NAR for validation. The NAR must verify whether the NCoA is acceptable, e.g. through DAD (Duplicate Address Detection) mechanism. In this paper, we assume the NCoA's validity is confirmed through DAD mechanism. If possible, the MN shall wait for FBack message on the PAR's link. If not, the MN shall resend a FBU as soon as it attaches to NAR. Depending on where the MN sends its FBU, in [2], two operation modes are defined: predictive and reactive. If the FBU is sent on PAR's link, the MN works in predictive operation mode, otherwise if it is sent on NAR's link, the MN works in reactive operation mode. In order to detail the working flow of FMIPv6, we classify the predictive mode as two sub modes: predictive mode I (FBack on previous link) and predictive mode II (FBack on new link). Note that this definition of operation mode is slightly different from that of [2].

After finishing link layer handover, the MN must send Fast Neighbor Advertisement (FNA) message immediately to NAR to inform it of the existence of MN in the new link. Whether releasing buffered packets right now or validating NCoA depends on the operation mode of the MN. In the following, the working flows of FMIPv6 are depicted according to MN's operation mode.

2.1 Predictive mode I (FBack on previous link)

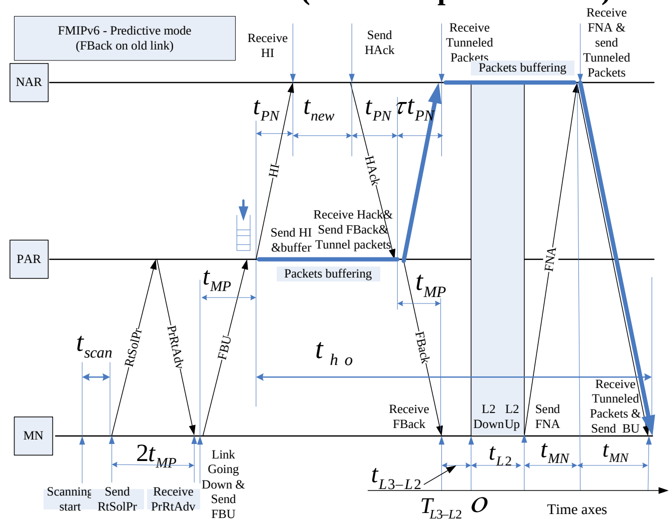


Fig. 1: Timing diagram of predictive mode I.

Figure 1 demonstrates the timing diagram of FMIPv6 in predictive mode I, in which the MN receives FBAck on the previous link. A similar timing diagram is presented in [3]. To make this paper self-contained, we re-draw it with different symbols.

t_{MP} : Packet transmission delay between MN and PAR

t_{MN} : Packet transmission delay between MN and NAR

t_{PN} : Packet transmission delay between PAR and NAR in wired network.

τt_{PN} : Transmission delay between PAR and NAR for tunneled packets. The τ is a weight for tunneling.

t_{new} : NCoA confirmation latency (for DAD).

t_{scan} : Scanning latency for 802.11 WIFI

t_{L3-L2} : Time interval between the FBack received on the MN in PAR's link and L2 link down in predictive mode I. In predictive mode II, it is the time interval between FBack which should be received in PAR's link if the link were not down and actual L2 link down.

T_{L3-L2} : The time when the FBack is received in predictive mode I with reference to the time origin. In predictive mode II, it's the time when the FBack should be received in PAR's link if the link were not down with reference to the time origin.

δ : Packet transmission latency per hop in wired network

t_{L2} : Link layer handover latency

For correct analysis, we set the time when the link is down as a time origin. In addition, we only consider downlink traffics. In figure 1, the PAR starts buffering for the MN on receipt of FBU. From then on, the MN cannot receive packets any more in the previous link. So we denote it the beginning of handover. After receiving the Hack message from NAR, PAR then forwards these buffered packets to NAR through the established tunnel. The NAR must intercept these tunneled packets and buffers them until it receives message FNA from the MN. The message FNA allows the NAR to forward the buffered packets to the MN. In [10], the handover latency is defined as the difference between the time a MN is last able to send and/or receive an IP packet by way of PAR, and the time the MN is able to send and /or receive an IP packet through the NAR. According to this definition, we get the expression of downlink handover latency.

$$t_{ho} = 2t_{PN} + t_{new} + t_{MP} + 2t_{MN} + t_{L2} + t_{L3-L2} \qquad (1)$$

We omit the packet loss during scanning period in the 802.11 WIFI network. Therefore, there is no packet loss in predictive mode I thanks to buffering mechanism. We define C_{loss}^{P1} as the number of packets lost during the handover process in downlink. Then

$$C^{P1}_{loss} = 0 \qquad (2)$$

And the required buffer size is given by:

$$C_{buff}^{P1} = \lambda(t_{ho} - \tau t_{PN} - t_{MN}) \qquad (3)$$

Symbol λ is the packet arrival rate.

We define packet cost as a weighted sum of required buffer size and packet loss given by (2) and (3), where w1 and w2 are weights.

$$C_{total}^{P1} = w_1 C_{buff}^{P1} + w_2 C_{loss}^{P1} \qquad (4)$$

2.2 Predictive mode II (FBack on new link)

The MN sends a FBU that includes the proposed NCoA at a time determined by LGD trigger from PAR's link whenever the "anticipation" of handover is feasible [2]. If the MN moves too fast to receive FBack on previous link, it should resend an FBU encapsulated in FNA immediately after link layer handover. The FBack will be sent from PAR to MN through wired network with the help of NAR. Figure 2 demonstrates this working flow in predictive mode II. When the MN operates in predictive mode I, the FBack is always received before the link layer handover, so T_{L3-L2} is smaller than zero in this case. However, in predictive mode II, the FBack can not be received by the MN in the previous link as link layer handover has already started. For analytical convenience, we still set the time when the FBack is received on the MN if the link is not down as T_{L3-L2}. In this case, the time T_{L3-L2} is positive. Therefore, when T_{L3-L2} is smaller than zero, the MN operates in predictive mode I. If it is bigger than zero and smaller than $2t_{MP} + 2t_{PN} + t_{new}$ (see reactive mode analysis) it means the MN moves too fast to receive FBack on previous link and the MN operates in predictive mode II. If T_{L3-L2} is still bigger than $2t_{MP} + 2t_{PN} + t_{new}$, this condition means there does not exits any link layer trigger and the reactive mode is issued. We denote this time threshold between predictive mode II and reactive mode as $T_2 = 2t_{MP} + 2t_{PN} + t_{new}$.

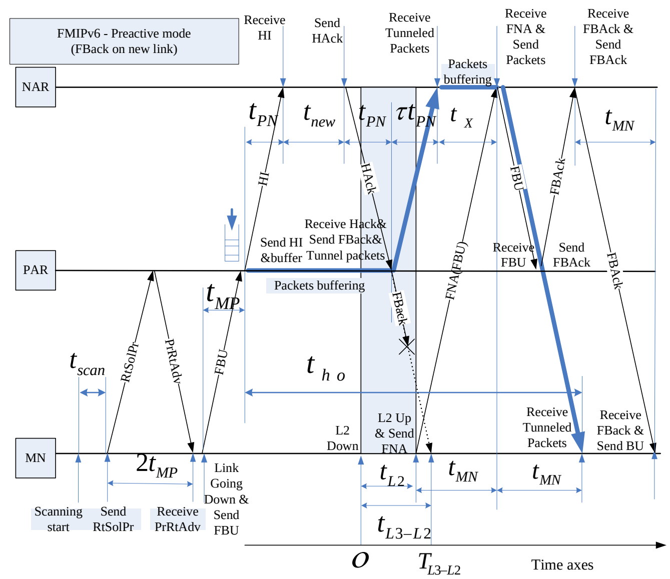


Fig. 2: Timing diagram of predictive mode II.

In the figure 2, the t_x is the time interval between the reception of the first packet received from PAR and the reception of FNA from the MN. On receipt of FNA, the NAR will forward packets tunneled from the PAR to the MN. The condition for t_x = 0 is that arrival time of FNA is smaller than that of receiving tunneled packets, that is

$$t_{L2} + t_{MN} \le t_{L3-L2} - t_{MP} + \tau t_{PN}$$

$$t_{L3-L2} \ge t_{MP} + t_{L2} + t_{MN} - \tau t_{PN} \qquad (5)$$

We define $T_1 = t_{MP} + t_{L2} + t_{MN} - \tau t_{PN}$

So in predictive mode II, (5) can be expressed as

$$t_x = 0 \quad \text{Whe } T_1 < T_{L3-L2} < T_2 \qquad (6)$$

If $0 < T_{L3-L2} \le T_1$, the t_x can be written as

$$t_x = t_{L2} + t_{MN} - (t_{L3-L2} - t_{MP} + \tau t_{PN})$$
$$= t_{L2} + t_{MN} - t_{L3-L2} + t_{MP} - \tau t_{PN}$$

In predictive mode II, the expression for downlink handover latency is given by

$$t_{ho} = (2+\tau)t_{PN} + t_{new} + t_x + t_{MN} \qquad (7)$$

The packet loss and required buffer size are given by the following expressions.

$$C^{P2}_{loss} = 0 \tag{8}$$

$$C^{P2}_{buff} = \lambda(t_{ho} - \tau t_{PN} - t_{MN}) = \lambda(2t_{PN} + t_{new} + t_x) \tag{9}$$

The total packet cost is written as

$$C^{P2}_{total} = w_1 C^{P2}_{buff} + w_2 C^{P2}_{loss} \tag{10}$$

2.3 Reactive mode

If the link layer trigger is not implemented, the MN enters into reactive mode after link up. In this case, the NCoA corresponding to the message FNA must be validated before using it. If the NCoA is acceptable, the NAR will forward those packets tunneled from PAR to the MN thereafter. But before the NCoA is accepted, there exists packet loss because the PAR has not received FBU. Therefore, the downlink handover latency begins from when the link handover is executed to when the first packet is received on new link, which is written as:

$$t_{ho} = (1+\tau)t_{PN} + t_{new} + 2t_{MN} + t_{L2} \tag{11}$$

And the resulting downlink packet loss and required buffer size are given as follows:

$$C^{R}_{loss} = \lambda(t_{L2} + t_{MN} + t_{new} + t_{PN}) \tag{12}$$

$$C^{R}_{buff} = 0 \tag{13}$$

The total packet cost is give by

$$C^{R}_{total} = w_1 C^{R}_{buff} + w_2 C^{R}_{loss} \tag{14}$$

Note that there is no FBU sent on the previous link in reactive mode, therefore the T_{L3-L2} shall not exist any more in this mode. But we may assume that if T_{L3-L2} is bigger than T_2, the MN operates in reactive mode.

3. The timing analysis for the TCP traffic

In the former section, the performance of the UDP traffic in FMIPv6 has been analyzed in detail. In this section, we will analyze the handover performance of a simplified TCP traffic on the basis of the following assumptions:

1) The data packets are sent from corresponding node (CN) to a MN. The initial congestion window of this TCP connection is one, and the slow start threshold is set as half of congestion window in steady state when retransmission timer expires.

2) In predictive mode, before PAR receives FBU and begins to buffer packets, the TCP connection operates in the steady state in the old network. If in reactive mode, the TCP connection also operates in the steady state before the link layer handover takes place.

3) All the packets received by the PAR before the FBU arrives in predictive mode, or before the link layer handover in reactive mode, have been properly Acked and all these ACKs are received by the CN.

4) The MN sends an ACK for every received data segment.

5) The DAD delay is much smaller than that of link layer handover by utilizing e.g. tentative address which an AR provides in link layer beacon frame[13] [9].

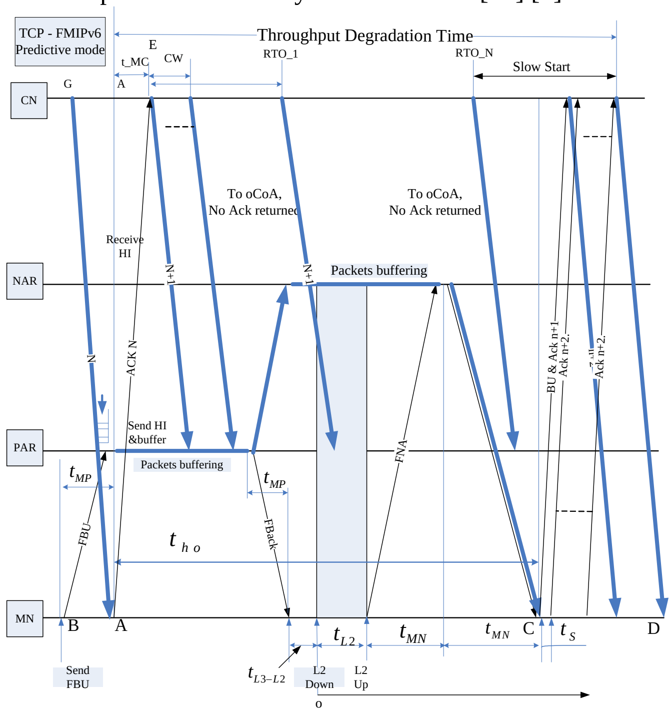


Fig. 3: Timing diagram of a TCP connection in predictive mode

In order to clearly specify the influence of FMIPv6 handover procedure to TCP traffics, we redraw the handover timing diagram of the predictive mode in figure 3. We denote the sequence number of the packet received at time A as N. Therefore, the MN is expecting the next packet with sequence number N+1. The last Ack of packet N is received by the CN at time E. Then, the CN transmits up to CW number of packets, where CW is its congestion window. If the CN does not receives the Acks for these transmitted packets at the time the first RTO occurs (denotes RTO_1), then the CN reduces the congestion window to one and retransmits these packets. At time C, the MN receives buffered packets from NAR and sends BU and Acks to CN to inform it of the new NCoA. On receipt of the Acks, the CN initiates slow start and ends it at time D.

In our scenario, the handover latency in the predictive mode generally is smaller than RTO if the signal threshold for link layer trigger is set properly and the delay for DAD is omitted. Therefore in figure 3, we can easily find that the downlink handover latency of TCP is equal to that of UDP traffic in figure 1 thanks to the buffer mechanism in PAR and NAR.

As to the reactive mode, the calculation of TCP downlink handover latency becomes rather complicated as a result of packet loss. They highly depend on the quantitive relations between RTO and downlink handover latency in reactive mode. In order to simplify the calculation, we also assume that there exits a tunnel between PAR and NAR for packet transferring in reactive mode and it remains active until a timer expires. This timer is a system design parameter and its period is bigger than TCP downlink handover latency assuredly.

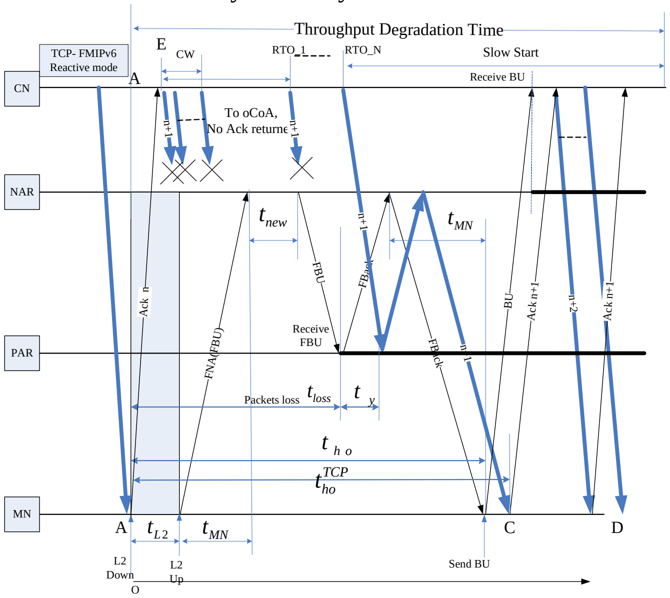


Fig. 4: Timing diagram of a TCP connection in reactive mode

In figure 4, a tunnel exits once the PAR receives the FBU message sent from the NAR and the packets destined to oCoA are forwarded to the NAR. During this period, the packet sent by the n-th retransmission is received by PAR and the MN receives it at the time C. But the packet of the (n-1)-th retransmission is lost due to the packet loss period t_{loss}.

So we get the downlink handover latency of a TCP connection in figure 4:

$$t_{ho}^{TCP} = t_{RTO}^{n} + t_{MC\_o} + t_{PC} + \tau t_{PN} + t_{MN} \tag{15}$$

$$t_{RTO}^{n} = (2^{n+1} - 1) RTO \tag{16}$$

Where the t_{RTO}^{n} is n-th TCP retransmission timeout duration [11]. If the CN does not receive the ACK for a packet before the expiry of the retransmission timer, TCP retransmits the packet and increases the RTO duration by a factor of 2. The t_{MC_o} is transmission delay between MN and CN in the old network, and t_{PC} is the transmission delay between PAR and CN.

4. Performance analysis

4.1 Handover latency and packet cost of the UDP traffic

From the above timing analysis, we can see that the timing of link layer trigger influences the operation mode of FMIPv6 directly. We carry out numerical simulations by MATLAB with the following parameters: $\lambda = 100$ packets/sec, $w_1 = w_2 = 0.5$, $\tau = 1.2$, the hops between PAR and NAR are 2 and 4 for intra-domain and inter-domain respectively. That is $t_{PN} = 2\delta$ (intra-domain) or $t_{PN} = 4\delta$ (inter-domain). We let two parameters T_{L3-L2} and δ changeable and investigate their influences to downlink handover latency and packet cost when the traffic is a UDP traffic.

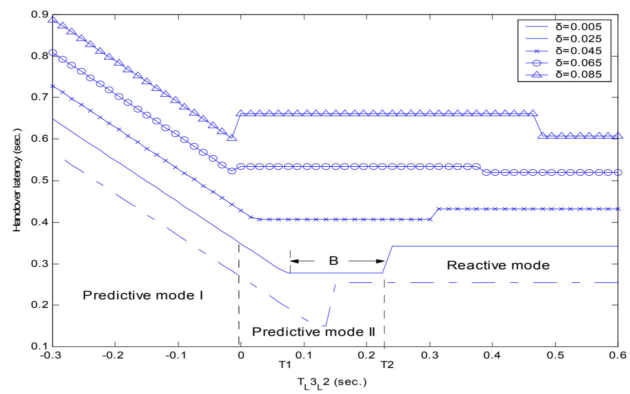


Fig. 5: Downlink handover latency with varying time T_{L3-L2}

Figure 5 explains the relation between downlink handover and time T_{L3-L2}. When T_{L3-L2} < 0, the downlink handover latency decreases linearly with T_{L3-L2}. It should be noted that when T_{L3-L2} is in the region B ([T1, T2]), the downlink handover latency is unchangeable, because in this predictive mode II t_x equals zero and the downlink handover latency is determined by tunneled packets' arrival time. While the MN is in reactive operation mode, the downlink handover has no relation with time T_{L3-L2}. Its downlink handover latency and corresponding packet cost is constant and determined given constant DAD time. Therefore, when δ is smaller than a certain threshold, region B is a desired work region for handover which can achieve minimum downlink handover latency and packet loss. The width of B is

$$B = T_2 - T_1 = (2+\tau)t_{PN} + t_{new} - t_{L2} \text{ When } t_{MP} = t_{MN} \quad (17)$$

Figure 6 specifies the relation between packet cost and time T_{L3-L2}. From figure 5 and figure 6, we can find out that when the transmission delay per hop between PAR and NAR increases, the width of region B also increases. The price for this benefit is the concurrent increase of downlink handover latency and packet cost. Note that when the transmission latency between PAR and NAR is beyond a certain threshold, the once best region B for handover reaches higher values than the reactive mode does. Based on the above analysis, we draw the following conclusions:

1) When transmission delay between PAR and NAR is below a certain threshold, in order to achieve best downlink handover performance, the FBU message is not necessarily received on the previous link.

2) When the MN operates in predictive mode, the downlink handover performance highly depends on the proper design of link layer trigger threshold. The additional anticipation time imposed by link triggers not only reduces the certainty about the MN's movement, but also increases downlink handover latency. Reference [6] proposed a QoS assuring trigger design method, but the complicated velocity estimation of the MN is needed.

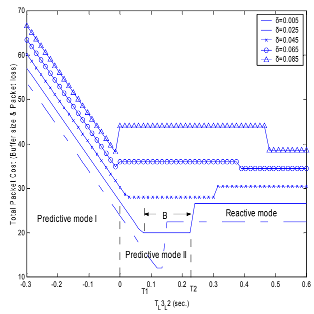
3) Although FMIPv6 achieves better handover performance than Mobile IPv6 does through advanced NCoA configuration and anticipated movement detection, the handover latency of link layer still is a limitation to it because of their mutual independences.

4) The downlink handover latency in reactive mode is always longer than the shortest one in predictive mode. However, its downlink handover latency is independent of link layer trigger time, and can be predicted, given network parameters such as transmission delay and DAD time.

5) This kind of mobile-controlled handover of FMIPv6 actually is called cell selection/reselection in 3GPP standards such as UMTS. In GERAN or UMTS, the PS (Packet Switch) handover process is controlled by the network. This kind of handover approach can provide improvements in terms of packets loss and handover delay.

4.2 Handover latency of the TCP traffic

In the simulation of TCP, we only care about metrics of the TCP handover latency, and the packet loss problem can be resolved by the TCP retransmission mechanism. Here, we let RTO = 0.5s, CW = 32, $t_{MC\_o} = t_{MC\_n} = 0.1s$, the other parameters are the same as those in the UDP simulation. In figure 7, the downlink handover latency of a TCP connection with the varying link layer trigger time is showed. This figure looks similar to figure 5, except it has higher downlink handover latencies in reactive mode than those of figure 5.



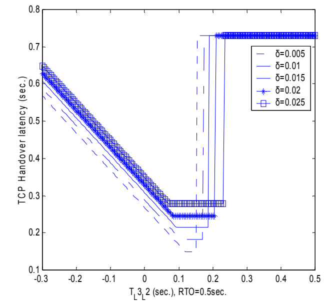


Fig. 6: (left) Packet Cost with varying time T_{L3-L2}.

Fig 7: (right) TCP Downlink Handover latency with varying trigger timer

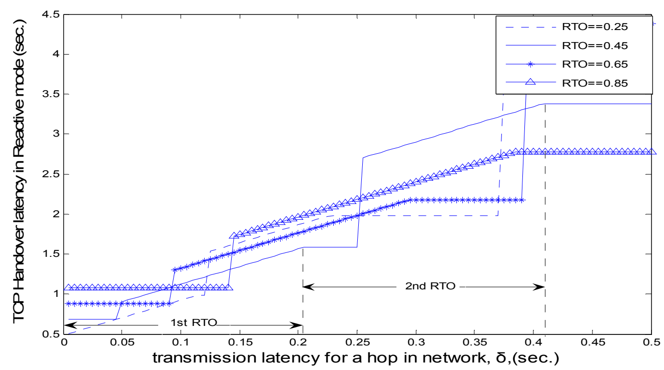


Fig 8: TCP Downlink Handover latency with varying transmission latency per hop between PAR and NAR in reactive mode

We also detail the downlink handover behavior in reactive mode with varying RTO. It can be seen from figure 8 that the TCP downlink handover latency is not linear as the transmission delay per hop between the PAR and NAR increases. In the case of RTO = 0.45, when the duration of the first retransmission timeout is bigger than the downlink handover latency of FMIPv6 in reactive mode, the handover latency of TCP is determined by the first retransmission of the packets. However, when the transmission delay per hop increase, it is likely that the first retransmitted packet is sent from PAR to NAR through an established tunnel and then to the MN. So, the TCP downlink handover latency is a linear function of transmission delay per hop. When transmission delay per hop exceeds a certain quantity, the first retransmitted packet is lost due to packet loss period of reactive mode. Under this condition, the handover latency of TCP is determined by the second retransmission. This is the reason why there exist two constant TCP downlink handover latency values at the beginnings of the 1st RTO and 2nd RTO in figure 8.

We conclude that the downlink handover latency of the TCP connection in FMIPv6 is very sensitive to packet loss of reactive mode. To avoid meaningless waiting time of retransmission timeout, it is expected that the FMIPv6 operates in predictive mode. Besides, the PAR needs to be sure of that the buffered packets are tunneled to the right target NAR to avoid packet loss. Therefore, we suggest that the network should take charge of the handover decision by providing the MN with the information of only one target network, such as a NAR's network prefix and L2 address after receiving the message RtSolPr in the WIFI. The MN shall respect the network's decision and make handover to the designated target network even it has scanned more than one AP.

5. Hybrid operation mode

Generally, the FMIPv6 is more suitable for reliable data traffics than MIPv6, but its downlink handover uncertainty is unwanted to: 1) Real-time traffic such as VoIP. For this kind of traffic, it often requires stringent handover delay, e.g. 40ms [7]. 2) Stringent handover latency scenario, e.g. 200ms for 802.20 [8].

In order to ensure downlink handover latency for the real time traffic, we propose a hybrid operation mode scheme that utilizes the certainty of reactive mode on downlink handover latency to satisfy these stringent latency demands. Before introducing the proposed scheme, we make the following assumptions: 1) The DAD delay is much smaller than that of link layer handover by utilizing e.g. tentative address which an AR provides in link layer beacon frame [9]. 2) The link layer handover latency is a system constant parameter to ARs, e.g. 50ms for 3G. Thus, when a predictive operation mode is triggered, the network calculates the prospective downlink handover latency of reactive mode. If it is below a delay threshold, the network chooses reactive mode by not sending back FBack message to the MN and continuously transmits packets on PAR's link. If not, the operation degrades to normal FMIPv6. The reason why the network controls the mode selection rather than the MN is in that the network dynamic state information such as transmission delay between ARs and the target network's RTD are only available to ARs. The hybrid operation mode consists of three parts as follows:

1) Measurements. The MN should periodically report measurement Round Trip Delay (RTD) to PAR in the message RtSolPr. The PAR receives the RTD measurement reports sent from many MNs and periodically averages them to get an estimation of dynamic network access time. In addition, the PAR should also periodically detect the transmission delay between itself and other neighbor ARs. During the handover execution, the current RTD of NAR's link should be sent back to PAR together with the verified NCoA in message Hack. The period for RTD reporting can be controlled by PAR by unsolicited message PrRtAdv if necessary.

2) Criteria. When the PAR received a FBU from a MN on its link, it carries out normal signaling exchanges with NAR. In Hack, the RTD of target network is returned. Then the PAR checks the traffic type destined to the MN. If it is real-time traffic, the PAR calculates the prospective downlink handover delay of reactive mode based on the following expression:

$$t_{ho} = (1+\tau)t_{PN} + 2t_{MN} + t_{L2} \tag{18}$$

If latency of (18) is smaller than predefined threshold, e.g. 200ms, the proposed hybrid scheme is initiated. Otherwise, the PAR executes conventional FMIPv6 predictive operation. Note that in (18), there is no latency for DAD compared with (11), since DAD time is much smaller than link layer handover latency by using pre-assigned tentative address [9], and in hybrid mode the delay for DAD can be omitted

3) Execution. If the hybrid mode is chosen, the PAR continuously transmits packets on its link without sending message FBack until it receives the message FBU from NAR. Hereafter, the packets destined to PCoA are forwarded to NAR through an established tunnel.

We carry out numerical simulation with MATLAB simulation tool for 2000 times for one MN. The handover mode are determined by the values of T_{L3-L2} which are uniformly distributed in the region [-0.3s, 0.3s]. The transmission delay per hop in network is changeable, while others are fixed and take the same values as before. If the downlink handover latency is above a threshold of 200ms, it is considered as a failed call.

From figure 9 we can find out the call drop rate of hybrid mode is much smaller than that of normal predictive mode under the condition that the transmission delay is below 0.015s. When the transmission delay goes beyond 0.015s, the handover operation regresses to the normal FMIPv6 performance. The hybrid mode has a delighted feature, in that the much bigger the DAD delay, the smaller call drop rate is. As a matter of fact when the DAD delay increases it influences the downlink handover of predictive mode instead of hybrid mode (again DAD time is much smaller than link layer handover latency). So the MN has more probability to choose hybrid mode.

Additionally, by inspecting carefully the curves of hybrid mode, we can see that call drop rate decreases linearly when transmission delay for one hop increases until 0.015s. In figure 5 we have stated the region B expands with the transmission delay for one hop grows. So it is more likely that predictive mode is triggered and more likely the hybrid mode is chosen subsequently given a fixed distribution region of T_{L3-L2}.

In figure 10, the average packet cost has a similar curve as call drop rate except the linear increase of cost in hybrid mode, because more hybrid selection means more packet loss.

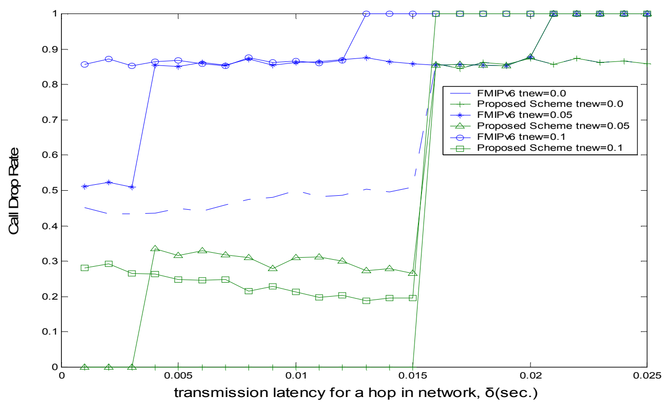


Fig. 9: Call drop rate with the variation of wired network transmission delay

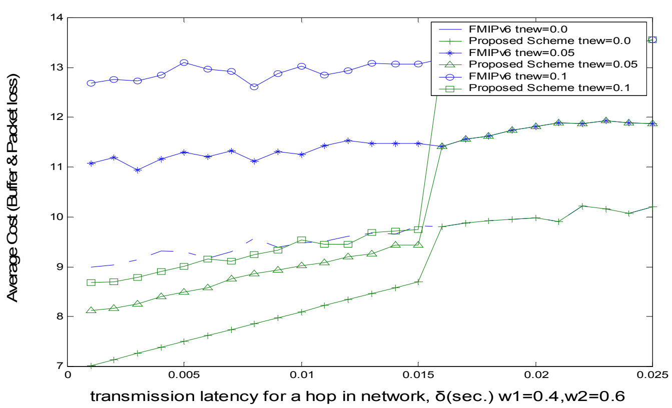


Fig. 10: Average packet cost with the variation of wired network transmission delay

6. Conclusion

The FMIPv6 reduces the long handover latency and high packet loss of MIPv6 by fast movement detection and fast binding update. But it suffers from uncertain additional anticipation time imposed by link layer trigger, especially for delay-constrained real-time traffic such as VoIP. We utilize the virtue of the certainty of downlink handover latency in reactive mode and pre-configuration of NCoA in predictive mode, to propose a hybrid mode for the real-time traffic which demands stringent handover delay but not low packet loss. The simulations show the hybrid mode can achieve lower call drop rate and lower packet cost when the network transmission delay is below a given threshold under the assumption that the DAD time is much smaller than link handover latency. In the case of throughput-sensitive traffics such as FTP, it is desirable for a TCP connection to operate in predictive mode with assured handover target. In order to reduce uncertainty of handover target, we suggest it is the network (PAR) makes the handover decision by providing the MN with target network information of only one NAR

7. References

[1] D. Johnson, C. Perkins, and J. Arkko. "IP Mobility Support in IPv6", http://www.ietf.org/rfc/rfc3775.txt, June 2004. IETF.

[2] R.Koodli. "Fast Handovers for Mobile IPv6", http://www.ietf.org/rfc/rfc4068.txt, July 2005. IETF.

[3] Seung_Hee Hwang, et.al. “Signaling Time Analysis for Optimal Fast Handovers for Mobile IPv6”, Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th, Volume 5, 26-29 Sept. 2004 Page(s):3275 - 3280 Vol. 5

[4] Lila Dimopoulou, et al, “Fast handoff support in a WLAN environment: challenges and perspectives”, IEEE Network, May/June 2005, pp14~20

[5] Emmanuel Seurre, et.al. “GPRS for Mobile Internet”, Artech House, 2003

[6] Shntidev Mohanty, et. Al.,”A cross-layer (layer 2+3) handoff Management Protocol for Next-Generation Wireless Systems”, IEEE transaction on mbile computing, vol.5, no.10,oct.2006

[7] Jon-Olov Vatn, “IP telephone: mobility and security”, Doctoral thesis, may 2005, KTH, Royal Institute of Technology, Stockholm, Sweden

[8] John L.Fan, “Desired Characteristics for an MBWA Air Interface”, IEEE C802.20-03/22

[9] Yoon, Sungro, et.al.” New Approach for Reducing DAD delay using Link Layer Assistance in Mobile IPv6”, Multimedia and Ubiquitous Engineering, 2007

[10] J.Manner, M. Kojo.,”Mobility Related Terminology”, RFC 3753, June 2004

[11] Shantidev Mohanty, Ian F.Akyildiz, “Performance analsis of Handoff Techniques Based on Mobile IP, TCP-Migrate, and SIP”, IEEE Trans. On Mobile Computing. Vol.6, No.7, july 2007

[12] N.Cardwell, S.Savage, “Modeling TCP Latency”, Proc.INFOCOM, Mar.2000

[13] V. Paxson, et.al. “Computing TCP’s Retransmission Timer”, RFC 2988, November 2000

[14] Yoon, Sungro, et.al.” New Approach for Reducing DAD delay using Link Layer Assistance in Mobile IPv6”, Multimedia and Ubiquitous Engineering, 2007