A Mobility Enhanced ATM Switch and Signalling Architecture to support handoff in a Wireless ATM Network

Farouk Smith, Neco Ventura
Dept. of Electrical Engineering, University of Cape Town
Rondebosch 7700, South Africa
{fsmith, neco}@crg.ee.uct.ac.za

Abstract—One of the most important aspects of mobility in a Wireless ATM (WATM) network is the ability to support a handoff. This paper describes a design for a Mobility Enhanced ATM Switch (MES) together with a handover signalling architecture to facilitate a fast and seamless handoff in a WATM network. The MES is simple enough to be implemented as a limited enhancement to a fixed ATM switch, yet provides low delay and lossless handover when used together with the handover signalling architecture. This paper reports on the experimental performance results of the MES and handover signalling architecture.

1. INTRODUCTION

The rise in wireless communications together with the increased need for higher bandwidth have signalled the start of a new era. Not only do users need high bandwidth, but they also demand mobility. With ATM identified as a viable broadband networking solution, Wireless Asynchronous Transfer Mode (WATM) has received increasing attention [1]. Therefore, it is reasonable to consider the extension of standard (fixed) ATM services into next generation wireless networks.

ATM was designed for environments where the hosts do not move, whereas with WATM mobility is a crucial factor. One very important feature of mobility is the ability to support the movement of a user between radio Access Points (APs), while maintaining unbroken communication. This movement between AP’s without a breakdown of the users connections is termed a handover and is required from the network.

The handover should be transparent to all the upper layers in the ATM user plane as well as the ATM adaptation layer and is defined as part of the ATM layer.

This paper describes a mechanism for a handoff protocol implementation as well as the design and implementation of a MES in order to facilitate a seamless handoff. This work is a continuation of the work presented in the analytical and simulation analysis of [18]. However, it’s focus is the proposal of a new two phase handoff scheme and the theoretical analysis thereof, whereas the focus of this paper is on the design of the MES and connection rerouting aspect of handover.

The paper is organised as follows. An overview of the WATM architecture is first presented with an explanation of handover schemes. Next, the objectives and design of the MES and handover protocols are discussed together with performance parameters associated with the handover protocols. The MES and handover signalling implementation is then described. Following this, the implementation results are given. Finally, some conclusions are drawn regarding the MES and handover signalling protocol implementation.

2. WIRELESS ATM OVERVIEW

In order to support mobility in ATM networks, the network infrastructure needs to have a set of network entities and functions [2]. A wireless access with ATM capabilities is needed to extend the same kind of transport service to the MT. An ATM signalling protocol is required that is capable of supporting connection establishment and seamless handoff when the mobile is migrating across boundaries, this would be termed the mobility signalling protocol. An Addressing and Location Management Scheme supporting both connectionless and connection-oriented traffic in a micro cellular network is needed as well as a wireless control mechanism for management of radio resources. Handoff delay must be very low and the handoff must be lossless. In order to avoid cell loss during the handoff, rerouting of the active connections and cell buffering are required in the network. Cell buffering and rerouting can cause cells to be out of sequence. The WATM architecture should provide for the efficient management of cell buffering and rerouting in order to achieve an in-sequence delivery of cells during a handoff. The quality of service of the connections during a handoff is affected by performance parameters such as the service disruption time during a handoff, time taken to complete a handoff and extra buffering needed to avoid data loss due to rerouting.

Fig. 1 WATM Network Architecture
Consider Fig. 1. There are two kinds of handoff, namely Intra-Switch and Inter-Switch Handoff. For Inter-Switch handoff the MT moves from one AP to another AP that is connected to a different switch as shown by process 2. Therefore, the handoff process requires a connection cross-point in the fixed ATM network. A Cross-over Switch (COS) is a connection break point on an original connection path in the network from which a new connection sub-path is established, i.e. the COS is a cross point of the old path and the new path. The COS is responsible for the actual internal switch rerouting of the connection. The COS is located in the fixed ATM network for an Inter-switch handoff, and hence, a rerouting of the connection is required in the fixed ATM network, and may not be an optimal (shortest) route. In Intra-switch handoff, the MT moves from one AP to another AP that is connected to the same switch, as shown in process 1. In this instance, the Mobility Enhanced ATM Switch (MES) acts as the COS, and is located in the WATM network. Hence, the new route is optional. The MES is a gateway between the fixed ATM network and the WATM network. The MES is a regular ATM switch with mobility control functionality. The WATM network is completely separated and independent from the fixed network. Therefore, the only access to the fixed ATM network for a Mobile Terminal (MT) is via the MES. The MT in turn is connected to the MES via an AP that is a gateway between the wireless and wired part of the WATM network.

Handoff Schemes for rerouting can be classified into four main groups. The full, partial, multipath and path extension re-establishment schemes.

The full re-establishment (FR) requires a completely new path, end-to-end, to be setup during handover [3]. The incremental re-establishment (PR) scheme requires only that a partial new path be setup and allows circuits to be re-used [4]. This scheme makes use of the COS. The Multicast re-establishment (MR) scheme makes use of multicast in ATM Networks [5], [6]. A controlling switch establishes a connection to the current serving AP and all AP’s in the neighborhood of the serving AP. When the MT moves to one of the neighboring AP’s, data is immediately available. Handoff is fast, but there is wastage of resources and this could result in call blocking for other MTs attempting to connect to the neighboring APs. In the path extension re-establishment (PE) scheme the route is extended from the original connection at the old MES to the destination connection at the new MES, where the new Access Point (AP) exist [7]. The PE scheme requires Permanent Virtual Circuits (PVCs) to be setup between adjacent MES to provide the path extension. After several handoffs, path looping occurs which results in a non-optimized path. The Two Phase handoff scheme has also been proposed that is a combination of the PR and PE schemes [8]. This technique is superior to the other handoff schemes because path extension reduces handoff delay and path (partial) rerouting increases resource utilization.

3. DESIGN OF THE MES AND SIGNALLING ARCHITECTURE

In order to accommodate handover in ATM switches, one needs to modify the switch controller to support mobility functions. The switch controller is a separate entity that controls the switch hardware. One of the functions to support mobility is the ability of an ATM switch to reroute connections dynamically while data is being transmitted through the switch fabric. The design of the MES should facilitate the rerouting of connections dynamically while data flows through the switch fabric.

With some modification and mobility support functions added to the switch control software, it is possible to implement the handover rerouting within the ATM switch at the ATM level. Before we present the design of the MES, we will examine the requirements of the MES in more detail.

As a whole, the MES should support the following requirements:

- The use of a mobility signalling protocol should be supported in order to facilitate a dynamic system. The signalling requests are conveyed to the system at the application level and not at the ATM or AAL level. The signalling messages will be responsible for the setting up and teardown of dynamic connections within the MES. The MES will also be able to route the signalling messages on separate VCs as the normal user data.
- Buffering is required by the system during a handoff because handoff in a WATM network introduces cell loss.
- The MES has to stop transmitting downlink cells destined for a mobile terminal immediately after it is identified as a COS.
- The system is required to provide in-sequence cell delivery during a handoff.
- The MES should be able to switch ATM cells dynamically while data is flowing to separate VCs in order to facilitate a handoff.

Fig. 2 illustrates the ATM cell flow within the MES [9]. The original connection path between the MT and its fixed terminal is routed through the switches input port 1 and output port 1. All the downlink cells (shaded in Fig. 2) from the host terminal are multiplexed with other ATM cells and arrive at input port 1 before handoff. These cells are routed to output port 1 to the current AP and hence to the MT. A handoff is then initiated when the MT moves to a new AP. The new AP is connected to the MES via output port j. After the handoff, all downlink cells needs to be sent to the MT in sequence and without loss. However, in a normal ATM switch, cells are routed within the switch fabric based on the VPI/VCI values of the ATM cells. When a handoff occurs, cells that are already stored in the output port buffer of the switch will be sent via the old route, to the old AP and be lost. Hence, some sort of buffering is required.

In order to solve this problem, one could, as indicated in Fig. 2, physically remove the cells in the output port buffer of the switch to some other buffer and route these cells to the input port buffer of the new connection. This is illustrated as process 1 in Fig. 2.

A new buffer for handoff purposes would have to be added to the switch hardware in order to avoid cell loss within the switch. Moving cells from one buffer to another also increases the end-to-end delay and wastes memory bandwidth. To realize these functions with a regular ATM switch is not straightforward and requires a hardware change. Our proposed system design of the MES requires no hardware additions to be made to the fixed ATM switch.

Instead of performing cell buffering within the ATM switch, the buffering is done elsewhere in order to simplify the process and to ensure rapid deployment of existing fixed
ATM switches without any hardware changes. Figure 3 illustrates the proposed system design of the MES with AP and MT buffers shown.

The original connection path between the MT and its fixed terminal is routed through the switch’s input port 1 and output port 1. All the downlink cells (shaded in Fig. 3) from the host terminal are multiplexed with other ATM cells and arrive at input port 1 before handoff. These cells are routed to output port 1 to the current AP and hence to the MT. A handoff is then initiated when the MT moves to a new AP. The new AP is connected to the MES via output port j. Once the ATM network has selected the MES as the COS (since the two APs are connected to the same switch), the following happens within the MES in order to ensure no cell loss and in-sequence delivery of cells:

- As soon as the MES receives confirmation that the link to the new AP is established, it forwards a last packet indication message to the MT via the old AP and reroutes the new downlink data via the new route.
- The MT will keep on receiving data via the old path until it receives the last packet and then starts to buffer data in the uplink direction.
- The MT will then try to acquire a channel to the new AP.
- The data in the new route for downlink direction is then buffered at the new AP until the MT acquires a channel to the new AP.

With this setup of external buffers, no changes are made to the switch hardware and any fixed ATM switch can be used with minor changes to the switch management functions in order to dynamically reroute connections.

The VXT entries in the COS are changed dynamically by the switch mobility control software that is located on an external control processor.

Only the PR and the PE schemes are considered in our design and implementation, as they are the basic building blocks of the Two Phase handoff scheme. These two schemes will be studied separately in order to determine their individual strengths and characteristics. With the characteristics, we explain why the Two Phase handoff scheme was proposed and what makes it superior to the other schemes.

The signalling sequence for the PR scheme for inter-switch handoff is shown in Figure 3. The signalling sequence for the PE is not given, but is done in a similar manner. Our proposed PR signalling sequence makes use of the Public Network Node Interface (PNNI) principle adopted for ATM networks and extended for WATM [10,11]. The Connection Server (CS) is a control switch that is a peer group leader of a collection of ATM switches and manages all switches in its group. Therefore, if a connection needs to be established in the group, the CS is contacted and computes an end-to-end route. If the connection spans many switches or administrative boundaries, the CS can in turn contact other CSs to establish partial routes to the desired destination. Rerouting during handoff only affects a small set of switches around the current serving AP and is handled by a single CS.

One important feature in this architecture is that the route information is centralized in a CS. This helps in identifying the COS immediately; the CS simply determines the shortest path between the new AP and the old route, and the point of intersection is the COS. If this particular route is heavily loaded or is unable to meet the QoS requirements for the connection, the CS will attempt to find another route. The CS then instructs the COS to reroute the connection.

The signalling sequence below is shown in Fig. 4:

1) When the MT detects the handoff occurrence in advance by the radio receive signal dropping below a set threshold value, it sends a handoff request (1) to the current serving AP, AP1.

2) AP1 then communicates to MES1 that a handover is needed by the MT with message (2). This message includes a list of the current active connections.
tions to the MT that need a handover to the new AP.
3) Immediately after receiving message (2), MES1 contacts the CS through messages (3) to locate the COS. The message includes the list of connections to be re-established.
4) Message (4) is a partial channel establishment from the COS to the new AP, AP2.
5) Message (5) is the acknowledgement of the partial channel establishment from AP2 to the COS. At this stage, the COS switches the downlink data and transmits via the new path. All data that is received by the new AP for the downlink is buffered until the MT acquires a channel to the new AP, message 11. The COS still receives uplink data via the old path. Hence, at this stage, only the downlink path is switched.
6) After the acknowledgement (5) the COS sends the last packet indication message for downlink data, messages (6), to the old AP, AP1 hop by hop along the old path.
7) Message (7) is the last packet indication in the wireless link from AP1 to the MT. After receiving message (7), the MT starts to buffer uplink data, and continues to do so until it acquires a channel to the new AP, message 11.
8) Message (8) is the handoff completion and connection release message from the MT to the old AP.
9) Message (9) is the handoff completion and connection release message from the old AP to the COS hop by hop. As soon as message (9) is received by the COS, it switches the uplink data VC to the new path.
10) Immediately after processing the last packet indication message (7), the MT acquires a channel to the new AP, AP2, via message (10).
11) Message (11) is the acknowledgement from AP2 to the MT. The time delay is denoted by $T_{11}$.

Based on the time taken by these signalling messages, various performance measurements can be derived [12]. The disruption time is the time interval between the instant the handoff completion command is received and the instant the first data packet is received by the MT in the new path. This includes the time to process all the signalling messages at the switches and the transmission time of the first data packet. It is the time for the MT to acquire a channel to the new AP plus the time it takes to receive a data packet from new AP.

The handoff completion time $T_{	ext{complete}}$ is the amount of time for all the rerouting to complete, i.e. from the time when the MT issues a handoff initiation request to the old AP to the time at which the connections to the old AP is torn down or the connection established to the new AP. All events except for the acknowledgement transmissions occur sequentially. Thus, the completion time is the sum of the times taken for each of the events during rerouting.

The amount of buffering required by the MT is determined by the amount of time during which the MT cannot transmit data on the wireless link. This includes the time for the MT to greet the new AP to acquire a channel and the time for the new AP to acknowledge the greeting.

The amount of buffering required on the new AP for the downlink is determined by the amount of data that is transmitted in the new path before the MT is connected to the new AP.

An analytical analysis was done to determine the above performance parameters based on the signalling messages and is reported in [13].

The signalling sequence for the PR scheme with a time perspective of the performance parameters is shown in Fig. 5. The X indicates the time at which the downlink and uplink data are switched.

![Fig. 5 PR signalling sequence with time perspective](image-url)

4. IMPLEMENTATION OF THE MES AND WATM HANDOVER ARCHITECTURE

The previous section described signalling concepts and the design of the MES. This section focuses on the implementation details of the MES and the handover protocols.

The evaluation platform for this work consists of the following hardware configuration. Five end stations, three serving as AP’s, one as the fixed host and the last one serving as the switch controller, are connected via multimode fiber to the Washington University Gigabit Switch (WUGS). The WUGS is a high speed, multicast virtual circuit experimental ATM switch developed at Washington University [14]. Although the WUGS was used for this experimental implementation, it is envisioned that the switch controller of any fixed ATM switch can be modified to support mobility.

Two WUGS switches are also connected via multimode fibers. The MT is connected to the AP’s via Ethernet to emulate the wireless access portion of the network.

There are no wireless ATM cards available for experimentation. However, this is not an obstacle, as in this investigation we are primarily interested in the effect of handover on the backbone ATM network portion of the connection. Hence, experiments can be conducted with the MT connected to the AP’s via Ethernet, Fig. 6. This would allow us to focus on the consequences of connection rerouting: loss, duplication, and reordering of packets.

The MES and WATM test bed is implemented on the Linux operating system. The Linux kernel is not ATM aware in its default state and therefore the ATM functionality has to be added before the kernel becomes ATM aware.

The protocol stack at the fixed host and the MT are different (Fig. 6). The fixed host uses AAL5 over ATM whereas the MT uses UDP/IP over Ethernet. The AP’s perform the bridging function between the different protocol
stacks. In an end-to-end ATM system, the mobile host would be responsible for reassembling the ATM cells.

The WUGS switch controller software consists of Jammer and GBNSC [14]. The WUGS control software resides on the Switch Controller PC as indicated in Fig. 6. The WUGS control software was modified to include mobility functionality. The application layer signalling will pass the handover request message to Jammer, which in turn pass the message to GBNSC. GBNSC then converts the signalling message into the ATM format that the switch understands.

Fig. 6 Experimental Test bed

At this stage it is worthwhile mentioning why the use of one or even two ATM switches in the experimental setup is sufficient in evaluating the performance of the various handoff schemes.

The effect of multiple switches on packet loss can be explained as follows with regards to the various handoff schemes. The period during which the translation tables in the COS are modified in order to reroute the connections, has the biggest impact on packet loss, [15]. Even when the connections span multiple ATM switches, the single COS at which the translation tables are altered will be mainly responsible for packet loss. The reason for this is that the handoff schemes first update all the switches in the new route and only then alter the entry at the COS. Therefore, it does not matter where in the connection the COS is located and the losses suffered due to multiple switch rerouting will be similar to our one hop experimental setup. However, in order to show the signalling delays when the distance between the MES and COS (Inter-switch Handoff) spans multiple switch, path looping within the same switch was performed.

We performed experiments by making use of continuous traffic, so that we can detect all potential disruptions during a handoff. A handoff request signal was generated with a manual interrupt signal at the MT.

5. RESULTS

By comparing log files of transmitted packets at the MT, AP and fixed host before and after a handoff occurred, it was found that data was transmitted without any loss and in sequence at all logging points. The MES was functionally correct on an end-to-end bases.

Fig. 7 shows the service disruption times of the Partial Re-Establishment and the Path Extension Schemes compared. The Path Extension scheme has a service disruption time similar to the Partial Re-Establishment Scheme. Both service disruption times are constant with an increase in the number of hops between the AP and the COS. The signalling sequence for both schemes are the same as far as the connection to the new AP is concerned, and the service disruption time is only affected by in the last hop of the handoff for both schemes and will not be affected by multiple switch handoff.

Fig. 7 Comparison of the service disruption time

The handoff completion times of the Partial Re-Establishment and Path Extension Schemes are compared in Fig. 8. The Path Extension scheme is much more superior to the Partial Re-Establishment Scheme as far as the handover completion times is concerned. Hence, handover network signalling is significantly reduced with the Path Extension Scheme. With the PR scheme, the new connection setup time is proportional to the number of switches between the COS and the MT. Therefore, as the number of hops between the COS and MT increase, so does the handover completion time. This is not an issue with the PE scheme, as the new path is simply extended to the new MES.

The buffer requirements at AP for the downlink are lowest for the PE scheme and are constant. The buffer requirements at the AP for the PR scheme increases with an increase in the number of hops between the COS and the MT.

Fig. 8 Comparison of the handoff completion time

The reason for this behavior is due to the fact that the amount of buffering required on the new AP for the downlink, is determined by the amount of data that is transmitted in the new path before the MT is connected to the new AP. This is directly related to the handoff completion time, since for the PR scheme the new connection setup time is proportional to the number of switches between the COS.
and the MT. Hence, more data will be transmitted in the new path the longer the handoff completion time.

![Fig. 9 Comparison of the buffer requirements at the AP](image1.png)

The buffer requirements at MT for uplink data during handoff are the same for both schemes, Fig. 10. This is because the amount of buffering required by the MT, is determined by the amount of time during which the MT cannot transmit data on the wireless link. This includes the time for the MT to greet the new AP to acquire a channel and the time for the new AP to acknowledge the greeting. The message sequence is the same for both schemes (messages 10 and 11, Fig. 4).

The results of this research correlates well with the analytical analysis of [3] and [9].

6. CONCLUSION

Handover is a very important function of a WATM network. The MES and handover signalling protocol design proposed in this paper is efficient and yet simple enough to be implemented as a limited enhancement to a fixed ATM switch. This would pave the way for a cost effective integration of fixed and WATM infrastructure. The relative simplicity of the proposed MES justifies the design objective that mobility could be added to fixed ATM switches, with very little modification and thus removing the need to make a distinction between switches for fixed and WATM networks.

We show that overall the PE scheme performs much better than the PR scheme and is in agreement with analytical predictions. However, the PE scheme is very inefficient when it comes to new route selection. Thus, by comparison of the PR and PE Schemes, one can easily see why the Two Phase Handoff Scheme was proposed in the literature. The Two Phase Handoff Scheme combines the PR and PE Schemes. The first phase of the Two Phase Handoff Scheme employs the PE Scheme to ensure a fast handoff delay and lower buffer requirements at the AP and MT. The second phase of the Two Phase Handoff Scheme involves a path optimization process, which is an implementation of the PR Scheme in order to optimize the new route.

10. REFERENCES