Design and Implementation of a functional WATM Test Bed to study the Performance of Handoff Schemes

Prepared by:
Farouk Smith

Supervised by:
Mr. M.J. Ventura

Department of Electrical Engineering
University of Cape Town
2003

This dissertation is submitted to the University of Cape Town in fulfillment of the academic requirements for the Degree of Master of Science in Electrical Engineering.

"There's no use trying," said Alice; "one can't believe impossible things."

"I daresay you haven't had much practice," said the Queen. "When I was younger, I always did it for half an hour a day. Why, sometimes I've believed as many as six impossible things before breakfast."

-- Lewis Carroll, "Alice in Wonderland"

Declaration

I declare that this thesis is my own work. Where collaboration with other people has taken place, or material generated by other researchers is included, the parties and/or material are indicated in the acknowledgements or references as appropriate.

This work is being submitted for the Master of Science Degree in Electrical Engineering at the University of Cape Town. It has not been submitted to any other university for any other degree or examination.

Farouk Smith | Date

Acknowledgements

I wish to thank the following people and organizations for their support in this research.

- Telkom, for their generous support and funding of this project.
- Mr. Neco Venture for supervising and reviewing my thesis.
- Dr. Jamal Elbergali, for his invaluable expert advice on WATM, his encouragement and also for reviewing my thesis.
- Mr. Sven Shepstone, for his advice and assistance with setting up the test bed
- My other friends and fellow students in the CRG and STAR group for their support and encouragement.
- My parents, for always believing in me and for their constant encouragement throughout my academic career.
- Last but no least, my wife and son, for their understanding, love and support during the two years that I have spent on this project.

Synopsis

Handoff is one of the most important functionalities required by a Wireless ATM (WATM) network to support mobility. The handoff is the process by which the user's radio link is transferred from one radio station (Access Point – AP) to another through the network without an interruption of the user connections.

WATM networks and fixed ATM networks are viewed as integral. Fixed ATM switches are thus enhanced with mobility specific functions that extend the services of the fixed ATM network to mobile users. These switches are termed Mobility Enhanced ATM Switches (MES) and are located at the boundary of the fixed ATM network. According to this network topology, 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 MES. Therefore, the handoff process requires a connection cross-point in the fixed ATM network. This connection cross-point is termed a Cross-over Switch (COS) and 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. Since the rerouting of the connection is required in the fixed ATM network, it may not be optimal. In Intra-switch handoff, the MT moves from one AP to another AP that is connected to the same MES. In this instance, the MES acts as the COS, located in the WATM network. Hence, the new route is optimal. The MES is a gateway between the fixed ATM network and the WATM network. 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.

To ensure a handoff without an interruption of the user connections, the handoff delay must be very low and handoff must be lossless. In order to avoid cell loss during the handoff, rerouting of the active connection and cell buffering are required in the network. Cell buffering and rerouting can cause a cell out of sequence problem. 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. One of the primary advantages of ATM is its ability to give QoS guarantees to connections. As mobile devices move between APs, QoS re-negotiation may be required to maintain levels of service to the connections.

The focus of this research is on the design and implementation of a WATM functional architecture in order to facilitate a seamless handoff. The project includes an experimental implementation of the WATM network. This required the building of a prototype WATM network with existing ATM switches and implementing handover protocol schemes at both the access and network sides.

It will be shown that a fixed ATM switch with minor modification, can be changed into a MES capable of supporting a fast handoff. 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 signaling architecture.

The handover rerouting schemes to be used in WATM networks were standardized by the ATM forum wireless workgroup in BTD-WATM-01.13. The handover signaling protocols are based on the Partial Re-Establishment (PR) and Path Extension (PE) schemes. We can measure the performance of the handoff schemes using the following metrics:

- The service disruption time, the time during which the MT cannot transmit data
- Handoff completion time, the total time for all the handoff signaling to complete
- Buffering required at the AP for buffering of downlink data during a handoff
- Buffering required at the MT for buffering of uplink data during a handoff

The values of theses metrics, evaluated for a given set of technology dependent parameters and various parameters of network connections, will give some idea of the effectiveness of the handoff schemes. An analytical analysis is developed in order to derive a mathematical model of the performance metrics and is later compared with the experimental implementation.

The experimental WATM implementation is used to evaluate the operational and performance metrics of the WATM functional architecture together with the handover signaling protocols. The experimental WATM network performs cell buffering at the APs and MTs and rerouting of the active connections in the MES. Mobility signaling is implemented at the MES, APs and MTs according to the applicable handoff scheme studied. With the experimental architecture, a series of operational and performance results of the WATM architecture and handover signaling protocols are presented.

The functional correctness of the WATM architecture on an end-to-end base was established 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. Results for the handover rerouting schemes shows that the PR scheme is the most optimal when it comes to new route selection, however, as the number of switch hops between the MES and COS increases for Inter-switch handoff, so does the handoff completion time. For the PE scheme it is found that the handoff completion time is always constant and fast compared to the PR scheme, however, the new route is not optimal for multiple switch handoffs. The service disruption time is constant and similar for both schemes and the PR scheme has a higher buffer requirement than the PE scheme at the APs and MTs. The results of the experimental implementation correlate well with the analytical analysis.

By comparison of the Partial Re-establishment and Path Extension Schemes, one can easily see why the Two Phase Handoff Scheme was proposed in the literature. The Two Phase Handoff Scheme combines the Partial Re-establishment and Path Extension Schemes. The first phase of the Two Phase Handoff Scheme employs the Path Extension 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 Partial Re-establishment Scheme in order to optimize the new route.

The experimental WATM architecture proposed in this thesis 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 infrastructures. 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 ATM and WATM networks.

Table of Contents

List of Figures

List of Tables

Glossary of Acronyms

ABR - Available Bit Rate

AP – Access Point

APDE - AP Data Entity

APSE - AP Signaling Entity

ATM - Asynchronous Transfer Mode

BER - Bit Error Rate

$BUF_{AP,down}$ - buffering required on the AP

BUF_{MT} - buffering required on the MT

BW_{wl} - Bandwidth of the wireless link

BW_{w} - Bandwidth of the wired backbone ATM network

CBR - Constant Bit Rate

CDMA - Code Division Multiple Access

CDV - Cell Delay Variation

CLR - Cell Loss Ratio

COS – Cross Over Switch

COSSE - COS Signaling Entity

CTD - Cell Transfer Delay

FHDE - Fixed Host Data Entity

FR – Full Re-establishment Scheme

GSM – Global System for Mobile telecommunications

HOC - Handover Control in the Network

HOC_{AP} - Handover Control in the AP

HOC_{COS} - Handover Control in the COS

HOC_{MES} - Handover Control in the MES

HOC_{MT} - Handover Control in the MT

H_{new} - Number of hops between the COS and the new AP

H_{old} - Number of hops between the COS and the old AP

H_{sig} - Number of hops between the old and new APs along the signaling channel

ISE - Intermediate Signaling Entity

LLC - Logical Link Control

L_{lw} - Latency of the wireless link

L_{w} - Latency of the wired backbone ATM network

LAN – Local Area Network

LPI – Last Packet Indicator

MAC – Medium Access Control

MES - Mobility Enhanced Switch

MESSE - MES Signaling Entity

MR – Multi-cast Re-Establishment Scheme

M-PNNI – Mobile Private Network Node Interface

M-UNI – Mobile User Node Interface

MT – Mobile Terminal

MTDE - MT Data Entity

MTSE - Mobile Terminal Signaling Entity

NTP - Network Time Protocol

PE – Path Extension Scheme

PHY - Physical Layer

PR – Partial Re-establishment Scheme

PCS – Personal Communication Services

PVC – Permanent Virtual Circuit

PPT_{fixed} - Protocol processing time for signaling messages

PPT_{adm} - Protocol processing time where admission control tests needs to be performed

RSSI – Receive Signal Strength Indicator

QoS – Quality of Service

RAL - Radio Access Layer

RMM - Radio Management Module

S_{sig} - Upper bound on the size of a signaling message

S_{data} - Maximum size of a data packet

TDD - Time Division Duplexing

TDMA - Time Division Multiple Access

$T_{acquire}$ - Time for the MT to acquire a wireless channel to the AP

$T_{complete}$ handover completion time.

$T_{disrupt}$ - service disruption time

UTC - Universal Coordinated Time

VC - Virtual Circuit

VCI - Virtual Channel Identifier

VPI - Virtual Path Identifier

WAN – Wide Area Network

WATM – Wireless ATM

WUGS - Washington University Gigabit Switch

Chapter 1

Introduction

1.1 Introduction

The rise in wireless communications together with the increased need for higher bandwidth has signaled 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, 2]. Therefore, it is reasonable to consider the extension of standard (fixed) ATM services into next generation wireless networks.

The harmonization of Personal Communication Services (PCS) and ATM, which have been developed somewhat independently, promises to be united by Wireless ATM into a single architectural framework. Mobile users will be able to carry high quality multimedia traffic.

To access backbone networks via WATM, users would request several types of virtual connections with agreed Quality of Service (QoS). User information may originate in any arbitrary format comprising continuous or variable bit rate voice, data image or video, and must fully utilize available bandwidth with multi-megabits per second access rates.

ATM was designed for very low bit error rates ($\sim 10^{-9}$), whereas wireless systems are typically designed to operate in bit error rates of between 10^{-3} to 10^{-6}. Wireless systems also have limited bandwidth resources compared to bandwidth rich ATM. With WATM mobility is a crucial factor that was not considered in the design of fixed ATM. One very important feature of a wireless network is the capability to support the movement of a user between radio Access Points (APs), while maintaining communication. This process is termed a handover and is required from the WATM network.

The aim of the handover protocol is to enable a wireless terminal to move between APs seamlessly while maintaining the negotiated QoS of its active connections. The efficient handover and switching of the Mobile Terminals (MTs) Virtual Connections (VCs) to the appropriate AP in the wireless network can achieve this. The methodology for implementing handover is one of the most important parts of WATM to support mobility [3], [4].

User requirements in ATM Networks demand multimedia applications with multiple active connections. These VCs will have different QoS requirements, and for distributed systems, the VC's will be connected to different end systems. During migration from one AP to another, the application program should be able to maintain the same connections as the terminal moves from one location to another. Also it should be possible to establish new connections to and from the MT regardless of its location within the network. Mobility should be supported by the network and user applications should function transparently during MT migration.

In order to support mobility in ATM networks, the network infrastructure needs to have a set of network entities and functions [1, 2]. A wireless access with ATM capabilities is needed to extend the same kind of transport service to the MT. An ATM signaling 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 signaling protocol. An Addressing and Location Management Scheme supporting both connectionless and connection oriented traffic in a micro cellular network is needed together with a wireless control mechanism to manage radio resources, handoff among APs and MTs. Handoff must be supported with low cell loss, latency and control overhead. The QoS guarantees must be maintained for each connection during and after MT migration, i.e. dynamic resource allocation, QoS provisioning and handoff control. Also, a MT can experience varying radio and network environments during migration, thus necessitating QoS renegotiation when the existing guarantees can no longer be met [5]. 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.
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 connection and cell buffering are required in the network. However, cell buffering and rerouting can cause a cell out of sequence

problem. 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 QoS of a connection during a handoff is affected by handoff performance metrics such as the service disruption time during handoff, time taken to complete the handoff and extra buffering needed to avoid data loss due to rerouting [6, 7, 8].

In order to integrate WATM into a fixed ATM network, two approaches can be adopted [9]. The WATM can be viewed as a completely separate network to the fixed ATM network, thus minimizing the impact on the fixed network. In this approach, mobility is mostly supported using separate network elements specific to mobility. A gateway would exist between the fixed ATM network and the WATM network and is the only link between the fixed ATM and wireless ATM network.

The other approach is to view the WATM and fixed ATM networks as integral. This view is also adopted by the ATM Forum specification BTD-WATM-01.13 [1]. Fixed ATM switches are enhanced with mobility specific functions. The resulting switch is termed a Mobility Enhanced ATM Switch (MES). Thus, the MES can support both fixed and wireless services, and no special types of switches are required. The MES should be 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 infrastructures. This approach will be adopted in this research and the network architecture is shown in Fig. 1.

The MES is a gateway between the fixed ATM network and the WATM network. The only access to the fixed ATM network for a 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.

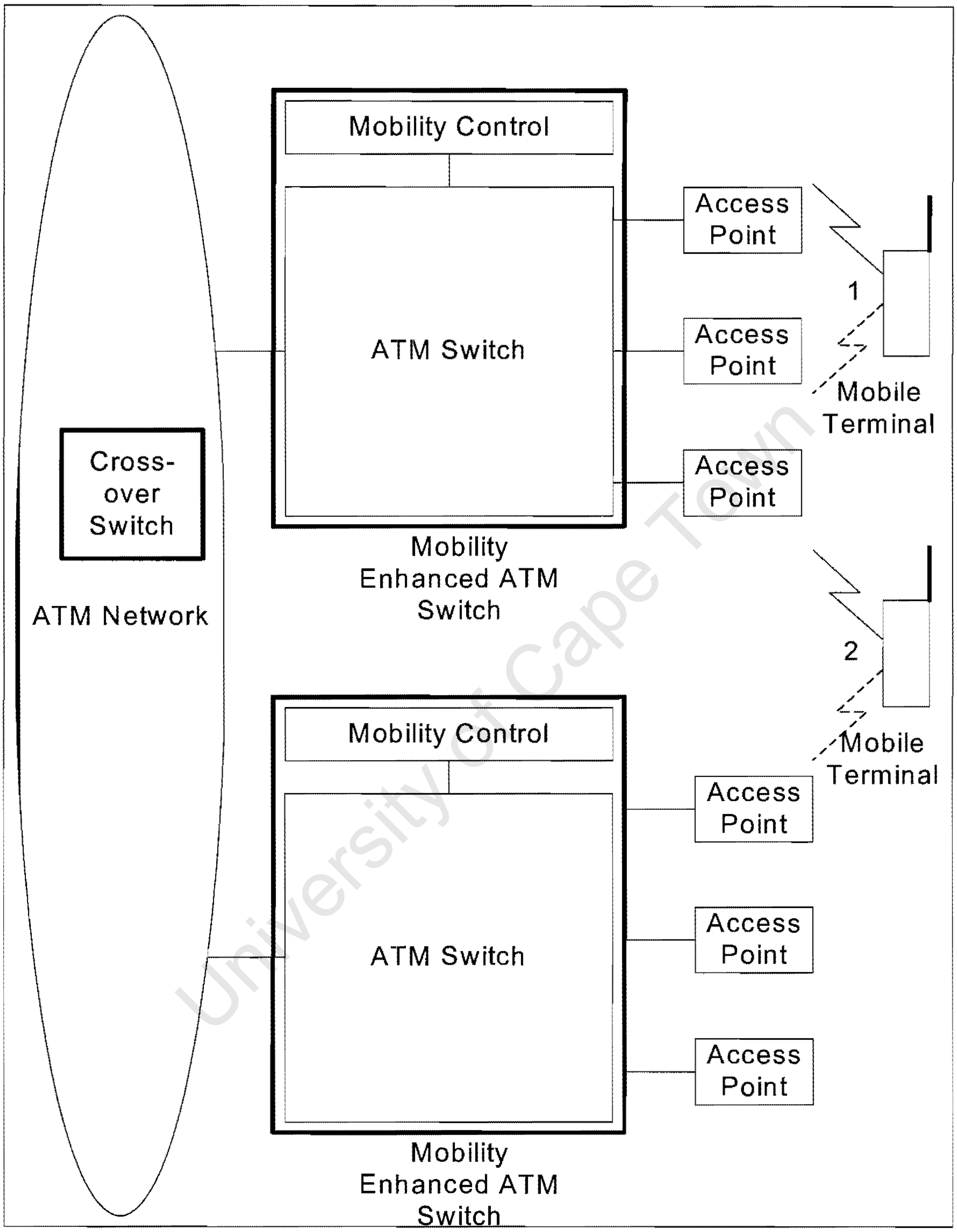


Fig. 1.1 WATM Network Architecture

The MES can be shared between fixed and wireless users and can be part of a larger ATM network. A separate protocol is used for the interactions of the MT and the mobility control functions in the MES. The AP receives ATM cells from the MES and translates them into a format suitable for wireless networks and sends them to the MT. On the other hand, the AP translates data received from the MT into ATM cells with the correct VPI/VCI and sends them to the fixed ATM backbone network via the MES. Thus, the AP controls the radio resources and is therefore responsible for the mapping of the VPI/VCI of an established radio channel to the MT.

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, process 2 in Fig. 1. 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 route. In Intra-switch handoff, the MT moves from one AP to another AP that is connected to the same switch, process 1 in Fig. 1. In this instance, the Mobility Enhanced ATM Switch (MES) acts as the COS. Hence, the new route is optimal.

Two functions needs to be supported to facilitate terminal mobility: connection handover, whereby an ATM connection from/to the MT is maintained by the wireless ATM network as the MT moves between APs; and location management, whereby the locations of MTs are tracked in the wireless ATM network to enable location-independent connection establishment to MTs.

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 to be setup during handover [8]. The Partial re-establishment (PR) scheme requires only that a partial new path be setup and allows circuits to be re-used [5, 6, 7, 8]. This scheme makes use of the COS. The Multicast re-establishment (MR) scheme makes use of multicast in ATM Networks [4, 10]. 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 a wastage of resources which 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 [11]. The PE scheme requires Permanent Virtual Circuits (PVCs) to be setup between adjacent MES to provide the path extension. For each handoff the path simply gets extended to a neighboring MES and after several handoffs could result in a non-optimized path. The Two Phase handoff scheme has also been proposed, i.e. a combination of the PR and PE schemes [12, 13]. This technique is superior to the other handoff schemes because path extension reduces handoff delay and path (partial) rerouting increases resource utilization.

The handover rerouting schemes to be used in WATM networks were standardised by the ATM forum wireless workgroup (BTD-WATM-01.13 [1]). The handover signaling protocols are based on the PR and PE schemes. However, it is specified that the two schemes be used as a combination to facilitate the Two Phase Handoff scheme.

As previously stated, the WATM functional architecture is defined in terms of various functional entities (e.g. MT, AP, MES and COS) and information flows (e.g. handover signaling protocols). The MT functionality is logically partitioned into ATM and radio layers. A similar partition is present in the network side also. The logical functions in the network can be implemented in different ways, depending on the placement of the physical realizations of the radio and ATM layer functions.

This thesis describes a method to physically design and implement the logical functions required by the WATM functional architecture in order to support handoff.

1.2 Thesis Objectives

The focus of this thesis is on the design and implementation of the WATM functional architecture in order to facilitate a seamless handoff. We therefore are not concerned with the location management functionality. The project includes an experimental implementation of the WATM network. This required the building of a prototype WATM network with existing ATM switches and implementing the handover protocol stacks at both the access and network sides.

This study also involves developing control-plane signaling modules at the various WATM functional entities to provide the necessary handover signaling functionality. The signaling entity will be integrated with the WATM user-plane functional components.

One of the WATM functional entities is the MES. In order to accommodate handover in fixed 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.

The MES should be simple enough to be implemented as a limited enhancement to a fixed ATM switch, yet provide low delay and lossless handover when used together with the handover signaling protocols.

The handoff signaling protocol is responsible for dynamically establishing new connections for handover, by means of signaling. Not only should the handover signaling protocol perform correctly, but it should also be efficient, in terms of handover signaling delays.

With the implementation of the functional WATM architecture, we evaluate the performance of the standardized handoff protocols, i.e. the PE and PR schemes.

Of particular interest will be to determine how the handover signaling protocols perform in terms of the following metrics: handover completion time, service disruption time during a handover, and buffer requirements at the AP and MT associated with handover and how the handover protocols could be optimised. Also of interest will be to determine how the handover signaling performance metrics mentioned above will be affected with an increase in the number of hops between the MES and COS for an Inter-switch handoff.

By making use of an analytical model of the handover protocols and a characterization of the WATM network, one can derive mathematical expressions for the metrics listed above. The work in this project also involves (apart from the experimental evaluation) an analysis of the PE and PR schemes.

With the performance metrics of the PR and PE schemes, it is shown how the characteristics of each can be used to facilitate the Two Phase Handoff scheme. The Two Phase Handoff scheme is a more efficient and a faster rerouting scheme than both the PR and PE if they are used in isolation.

1.3 Scope and Limitations

The scope of this study is limited to developing a functional user and control plane WATM test bed to facilitate a handoff. In particular, the MES design is implemented on the Washington University Gigabit Switch (WUGS) and integrated with the handover signaling protocols. The WUGS is a high speed, multicast virtual circuit experimental ATM switch developed at Washington University. The design of the MES was based on modifying an existing fixed ATM switch (WUGS) to support mobility. Even though many techniques could be employed to enhance the performance of the MES and handover protocols, this study could not address every issue that could impact the performance. The development of the user and control plane signaling entities are highly specialised and no commercial tools are available to aid in implementing a suitable design. As wireless ATM is in its infant stages, no MES or mobility control plane signaling architecture is available commercially to the author's knowledge.

Any signaling protocol is unable to establish an end-to-end user connection without the existence of a previously established ATM VC; this is termed the signaling VC. Hence, the signaling protocol cannot create a VC to support itself. The same is applicable with mobility signaling. It is impossible for a MT to request a handover from the network without an already existing underlying ATM VC, which is reserved for signaling. The handover signaling protocol therefore requires that an existing VC exists between the various network elements that requires mobility signaling. The mobility signaling protocol implementation is completely separate from fixed ATM signaling and is thus unable to request an increase in ATM VC bandwidth from the fixed ATM signaling protocol if, after handoff, the new connection does not satisfy the previous QoS of the connection. We assume and specify that the new connection **does** satisfy the QoS of the previous connection, therefore not concentrating on dynamically allocating bandwidth.

Even though the user and control plane signaling entities were designed to optimize the handover performance metrics, the physical hardware utilized in the implementation provided some limitations. The evaluation platform for this work consists of the following hardware configuration. Five end stations, two serving as AP's, one as the fixed host, one serving as the switch controller and the last one as the MT. All end stations except the MT are connected via multimode fiber to the WUGS. Although the WUGS was used for this experimental implementation, the switch controller of any fixed ATM switch can be modified in a similar manner to support mobility if the switch control software has a publicly available source code. 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. This would allow us to focus on the consequences of connection rerouting: loss, duplication, and reordering of packets.

The use of one or even two ATM switches (only two WUGS switches are available for research purposes) in the experimental setup is sufficient in evaluating the performance of the various handoff schemes. The effect of multiple switches on packet loss with regards to the various handoff schemes can be explained as follows:

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. 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 switch experimental setup. However, in order to show the signaling delays when the distance between the MES and COS (Inter-switch Handoff) spans multiple switches, path looping within the same switch was performed.

Experiments can be performed with audio and low bit rate video as user traffic, in order to quantify the service disruption time, and buffer requirements at the APs and MT due to the handover. Audio traffic can be characterized by its timing component. Every T milliseconds, a 160-byte packet will be transmitted by the source. For T = 20, this corresponds to telephone quality audio of 64 kb/s. In a typical audio conversation, there will be silence periods during which no packets will be transmitted. If this occurs during a handoff, there would be no disruption of traffic. It is for this reason that we will make use of continuous traffic, so that we can detect all potential disruptions during a handoff. Low bit rate video traffic can be characterized by a Constant Bit Rate source sending packets every 30 ms. This would correspond to a frame rate of approximately 30 frames per second, the frame rate for full motion video. Therefore, instead of transmitting real audio and low bit rate video, we performed experiments by making use of a continuous traffic (Constant Bit Rate – CBR) source at the MT and fixed host, so that we can detect all potential disruptions during a handoff. A handoff request signal at the MT was generated with a manual interrupt signal at the MT.

It is strongly noted that we are only concerned with the design of the physical user and control plane functional entities of a WATM network to facilitate a handoff. The thesis will also not concern itself with Crossover Switch (COS) discovery algorithms, but will rather allocate the position of the COS in the network statically, and it is therefore assumed that the network will have advance knowledge of where the COS is located.

We are also not concerned with the actual contents of the signaling messages; rather focusing our attention on the signaling sequence during a handoff.

1.4 Thesis Outline

This introduction forms the first chapter of this thesis. The remainder of the document is organized as follows:

Chapter 2: Background Theory and Literature Review. Chapter 2 presents an overall view of the important issues regarding handover in a WATM network as well as related research. The functional WATM architecture is first presented where various components are identified that will be important to facilitate the handover control platform. Various handoff rerouting schemes are then classified together with their features, advantages and disadvantages. Key features and assumptions influencing each scheme are also presented. The chapter then discusses issues related to cell routing and loss in the MES during a handover operation as well as performance metrics that will influence the quality of service of connections during a handover. The chapter concludes with a summary of the key issues regarding handover.

Chapter 3: Handover Signaling Analysis. Chapter 3 presents the signaling sequence of the PE and PR schemes together with an analysis. The analysis is performed by considering the signaling messages required to be exchanged during handoff in a common WATM network structure. A quantitative comparison of the PE and PR schemes is presented using analytically derived formulas for the performance metrics such as the: (i) Service disruption time, (ii) Handoff completion time, (iii) Buffer requirements at MT, and (iv) Buffer requirements at AP required during a handoff. With the analysis we show how the PE and PR schemes can be combined to form the Two Phase handoff scheme.

Chapter 4: System Design of the Handover Signaling Entities. This chapter begins with a description of the issues at hand with the design of the handover signaling entities. Several entities in the WATM network are identified to support the handoff. Chapter 4 presents the design and architecture of the MT Signaling Entity, the AP Signaling Entity, MES Signaling Entity and the COS Signaling entity. Further, chapter 4 explains that the framework developed should be based on a modular design methodology ensuring that future research employs this framework with minimal modification effort. This is achieved by selecting a modular design approach in developing the control plane handover signaling framework.

Chapter 5: WATM and Signaling Framework. This chapter focuses on the implementation details of the experimental WATM functional architecture and the handover protocols. It begins by presenting the architecture of the evaluation platform together with a description of its components. The method used to capture time and sequence numbers are then described, and following it a description of the implementation details regarding the user plane data entities. The chapter then describes the different experimental configurations considered and what method were used to emulate multiple hops between the MES and COS. The chapter concludes with a discussion of the user traffic sources used in the evaluation platform.

Chapter 6: Framework Evaluation and Results. This chapter discusses the evaluation of the WATM architecture and handover signaling protocols. The first section discusses the performance evaluation of the MES, where particular focus is placed on the MES ability to switch packets correctly during a handover. The rest of this chapter is devoted to discussing the results of the tests performed to evaluate the operational aspects of the WATM architecture and handover signaling protocols. The performance measurements to be considered with the experimental setup for the handover schemes are (i) Service disruption time, (ii) Handoff completion time, (iii) Buffer requirements at MT, and (iv) Buffer requirements at AP.

Chapter 7: Conclusions and Recommendations. This chapter presents a set of conclusions that were drawn from this study as well as recommendations that may be used for future work regarding WATM.

Appendices: Sets of appendices are presented to provide further background information.

1.5 Published Work

The design, implementation and results presented in this thesis were published in the proceeding of the South African Telecommunications and Networks Applications Conference (SATNAC) 2002 and 2003.

1) F. Smith, N. Ventura, "Evaluating the Performance of Handoff Schemes in Wireless ATM Networks", SATNAC 2002.
2) F. Smith, N. Ventura, "A Mobility Enhanced ATM Switch and Signaling Architecture to support handoff in a Wireless ATM Network", SATNAC 2003.

Chapter 2

Background Theory and Literature Review

2.1 Introduction

An important element of WATM is developing a MES that is capable to support a fast and seamless handover signaling architecture. The MES should be compatible with the principles of ATM cell relay and support the QoS requirements of both loss and delay sensitive multimedia services. By viewing WATM as an integral part of the fixed ATM network, as we do in this thesis, adds to this requirement to implement handover in a manner that is compatible with fixed ATM switches [9, 14]. Thus fixed ATM switches are enhanced with mobility functionality.

A very important constraint in our design is that the MES should require no changes to fixed ATM switching hardware. It should be feasible to use an existing ATM switch to support wireless and wired services without major changes to the switch. In addition to the MES, the AP, MT and COS functional entities also needs to be realized in order to facilitate a handoff.

Before a design of any network element can be addressed, one needs a thorough understanding of all the important issues involved. Therefore, some of the technical challenges associated with the handover in a WATM network that need to be addressed, are discussed in the sections that follow.

2.2 WATM Functional Architecture

The WATM functional architecture is defined in terms of various functional entities, information flows and physical reference points [1]. This is shown in Fig. 2.1. Here, the MT functionality is logically partitioned into ATM and radio layers. A similar partition is shown in the network side also. The logical functions in the network can be implemented in different ways, depending on the placement of the physical realizations of the radio and ATM layer functions. The network entity where the ATM layer interactions with the MT are first handled is the Radio Access Point (AP). The interfaces present in the WATM functional architecture are described below. However, this thesis only concentrates on the interfaces that are required to support a handoff in a WATM network.

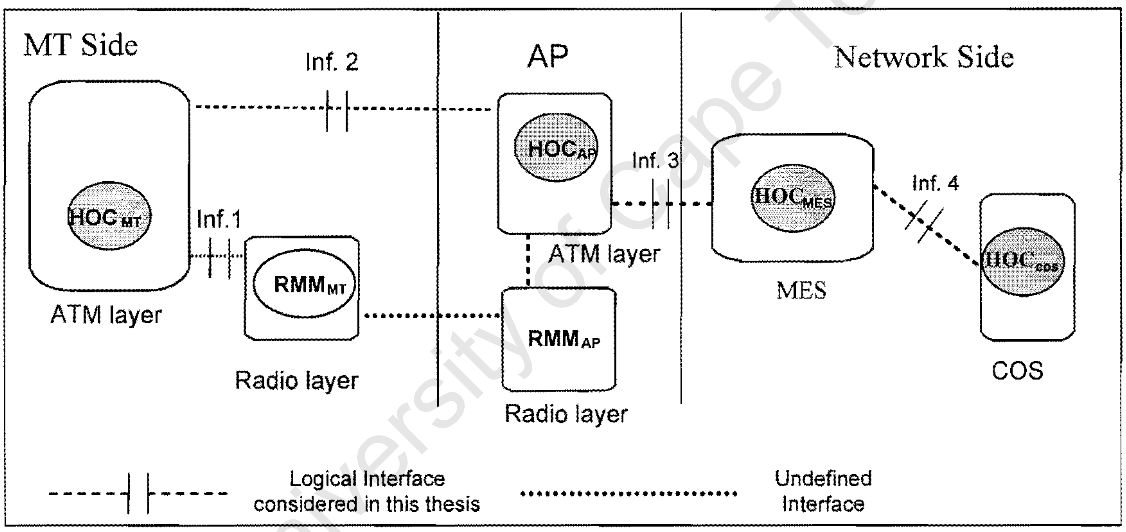


Fig. 2.1 WATM Functional Architecture

- Inf.1: This is the interface between the ATM and radio layers at the MT. This is a generic interface that must be supported by all WATM compatible radio layers. This thesis only considers the requirements to be satisfied at Inf.1, and do not concentrate on the details of its implementation.
- Inf.2: This is the interface between the MT and AP required to exchange handover signaling information.
- Inf.3: This represents the handover signaling interaction between MESs in the network and APs.

- Inf.4: This is the handover signaling between MESs and the COS (fixed network).

The interface between the ATM and radio layer in the network side is beyond the scope of this thesis. Furthermore, the interface at the radio layer between the MT and the AP is also not within the scope of this thesis.

2.3 WATM Functional Entities

This section describes the functional entities illustrated in Fig. 2.1.

2.3.1 HOC_{MT} (Handover Control in the MT)

This functional entity is responsible for:

- Deciding when active connections must be handed over to a new network point of attachment (AP);
- Initiating ATM-layer handover for active connections;
- Executing the ATM-layer handover protocols; and
- Interacting with higher layer applications when handover fails.

This is a crucial entity to support the connection rerouting aspect of the handover at the MT. The control plane module for this entity will be designed in chapter 4. This entity will later be referred to as the Mobile Terminal Signaling Entity (MTSE).

2.3.2 HOC (Handover Control in the Network)

This functional entity performs the complimentary functions related to handover as performed by the MT at the network side. This includes:

- Receiving handover requests from the MT;
- Deciding on which network point of attachment the connections must be handed over to;
- Executing the network side handover procedures; and
- Indicating handover results to the MT.

These entities consist of the HOC_{AP}, HOC_{MES} and the HOC_{COS}. They are crucial entities to support the connection rerouting aspect of the handover in the network side, and the control plane modules for these entities will be designed in chapter 4.

2.3.3 Functions covered in this thesis

While the functional model in Figure 2.1 describes the typical functions in the MT and the network to support a handover, this thesis does not cover all functions required to support terminal mobility. Specifically, the security, location and service management features are not part of this thesis. Call control, signaling and routing in the fixed ATM network is covered by fixed network ATM specifications [51 – 53] and this thesis describes only the incremental changes necessary to support handover. Thus, handover control and signaling are fully described in this thesis.

The Radio Management Module (RMM) is responsible for managing the physical radio resources and is also beyond the scope of this thesis. The MTSE only needs access to the radio layer in order to monitor its quality.

2.4 WATM Protocol Layering

Fig. 2.2 depicts the protocol layers required to support the WATM reference configuration shown in Figure 1.1. The following entities are seen:

- **M-PNNI:** PNNI with supplemental signaling information to support mobility. M-PNNI layer also supports pure PNNI functions, i.e., if the mobility features cannot be used then the device with the M-PNNI features can operate with the PNNI feature set.
- **M-UNI:** UNI with supplemental control signaling to support mobility.
- **UNI**: Regular UNI.
- **RAL:** Radio Access Layer. This realizes the wireless segment of the end-to-end ATM connection. It includes the wireless MAC, LLC, and PHY layers. In this thesis we only specify that the **HOC_{MT} (Handover Control in the MT)** should have access to the RAL in order to monitor the quality of the radio link. If the radio signaling strength drops below a set threshold value, the **HOC_{MT}** will initiate a handover operation.

- **ATM-CL**: The ATM Convergence Layer above RAL to transport ATM data over a generic RAL that provides the standard interface, Inf.1.
- **ATM**: The ATM layer.
- **PHY**: ATM physical layer.

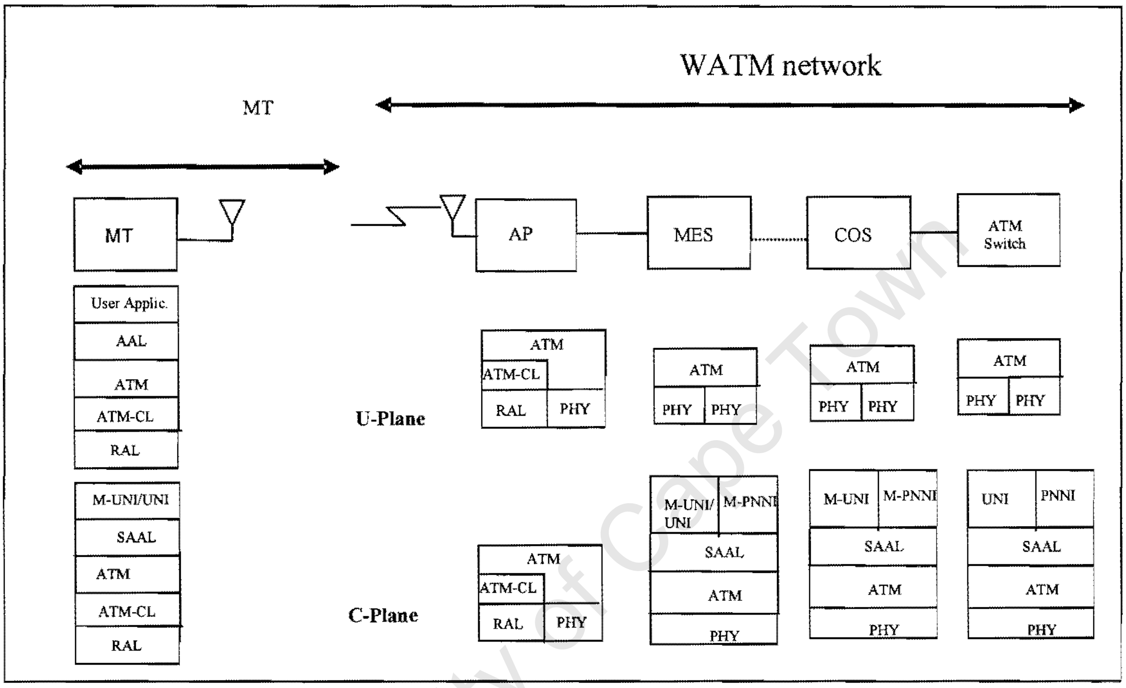


Fig. 2.2 WATM Protocol Layering

2.5 Rerouting Schemes for Handoff

In chapter 1 rerouting schemes for handoff were presented. The first in the group was classified as the Full Re-establishment (FR) scheme [8]. This scheme requires a completely new path to be setup between both end stations (MT and fixed host) during handover resulting in an optimal path. Since a completely new path is created end to end, it introduces large handover delays. This scheme is very inefficient and is referred to in the literature as a base for comparison to other handoff rerouting schemes.

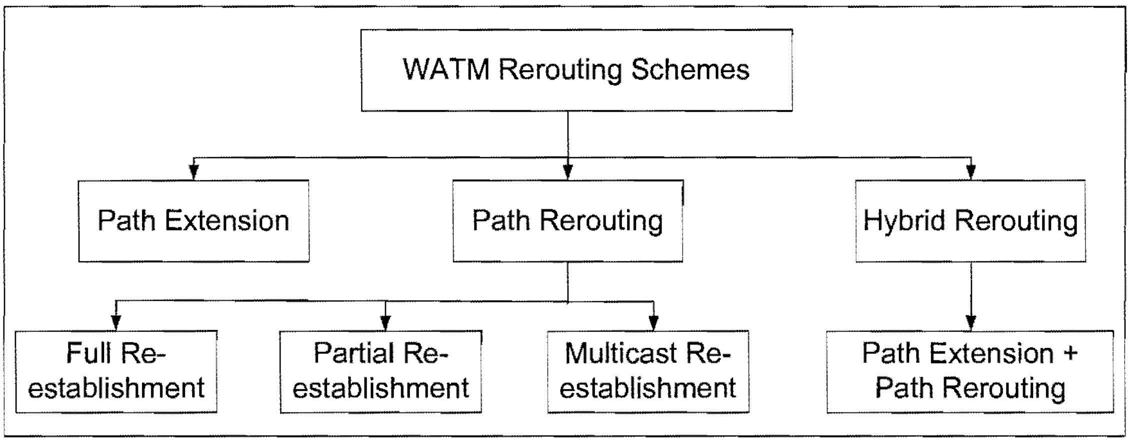


Fig. 2.3 Classification of Rerouting Schemes

In the Multicast Re-establishment Scheme (MR) [4, 10], the concept of a virtual connection tree concept was proposed. When a MT establishes a connection to the WATM network, a connection tree is formed from a root (an ATM switch) to a set of APs to which the MT may move. Whenever the MT moves into one of those APs, it uses one of the pre-established virtual channels from the root node to the corresponding AP. The goal of a virtual connection tree is to reduce handoff latency by removing connection setup time for the new connection sub-path. However, this method uses pre-established multiple virtual channels even though only one of them is used at a time. This requires more network resources (virtual channel space, bandwidth, and memory for routing information) than other schemes.

The path extension scheme extends an original connection path and adds a new connection sub-path from an old MES to a new MES which is connected to a new AP that the MT is going to visit [11]. This method is simple, and makes it easy to maintain ATM cell sequence during a handoff. However, for multiple handoffs the extended connection path will increase end-to-end latency.

In the Partial Re-Establishment Scheme (PR) [5 - 8], an original connection path between the MTs peer and an old base station is torn down and part of it is replaced with a new connection segment between a new base station and an intermediate switch on the original connection path during a handoff. Here we define a Cross-over Switch (COS) as a switch on the original connection path from which a new connection sub-path is established.

In a handoff procedure, a COS can be fixed [4, 18] or can be determined dynamically [19, 20]. The algorithms for selection of a COS determine the tradeoff between end-to-end routing efficiency and handoff latency. For an optimal end-to-end route, a complicated algorithm is needed for selection of a COS. This will increase handoff latency [20]. Fixed COS allocation is also known as static COS allocation. By statically allocating a COS, connections to the MT need not use a COS selection algorithm, rather connections are switched at a dedicated switch for the MT. The static COS is typically assigned per MT at the time of call admission. This simplifies the selection procedure, but does not result in optimal paths.

The Two Phase handoff scheme (Hybrid rerouting) has also been proposed, i.e. a combination of the PR and PE schemes [12, 13]. This technique is superior to the other handoff schemes because path extension reduces handoff delay and path (partial) rerouting increases resource utilization.

2.6 Handover Classes

WATM supports the following classes of handover: hard and soft, backward and forward handover [21].

- There are two kinds of rerouting for handover based on the number of active connections to the APs.
    - Hard Rerouting - The old AP is dropped before the new AP is added. Thus the MT has a radio connection with only one AP at any time.
    - Soft Rerouting – The new AP is added before the old AP is dropped. Hence, the MT is connected to two AP's simultaneously.
- Based on the direction of the handoff signaling, handoff can be classified in two ways:
    - Backward Handover – the old AP initiates the handover signaling after receiving a handover request from the MT. The MT would typically notice that the radio signal to the old AP is fading and therefore request that the old AP initiate a handoff. The MT

continues to transmit and receive data via the old AP during handover and only switches to the new AP after radio resources have been reserved and all entities involved are prepared for the handover.

- Forward Handover – A forward handover is characterized by a MT arriving at a new AP suddenly. The MT associates itself with the new AP and requests a handover via the new AP. The new AP then initiates the handover. This normally happens when the MT suddenly loses connection to the old AP and there is no time to initiate a backward handover.

These types of handover can be flexibly chosen, depending on radio conditions and QoS requirements, to enhance handover performance and robustness. The two basic handover mechanisms are hard and soft handover. In the case of hard handover, the handover control flow can be directed either across the current AP's air interface (backward handover) or across the target AP's one (forward handover) [23]. A number of handover protocols have been proposed based on hard handover [19, 22, 24]. In the case of having the radio link deteriorating rapidly, what is important is to have reliability and robustness of the handover signaling protocol. However, under these circumstances the regular backward handover mechanism will not perform reliably any more. In practice, the quality of the radio link prevents any communication on the dying radio link before the handover procedure can be completed. The signaling control flow will most probably be severely impacted [23]. Therefore the handover protocol has to be supplemented by a forward handover procedure, so as to avoid having handover failure and subsequent call dropping. The forward handover aim is to maintain a connection despite the fact of having unexpected link failures during a handover situation. However, this is achieved at the expense of losing data on the old mobile connection segment and a higher cost of resynchronization.

The advantage of having a backward handover mechanism is that it gives one the possibility to choose the best AP for the MT to connect to. In case of forward handover, the MT suddenly interrupts the old connection and tries to connect to a new AP. However, this new AP cannot reserve resources in advance and therefore may not be such a good solution. But if the current connection is interrupted by radio interference, this is the solution of reconnecting fast enough.

A number of handover schemes have been designed in such a way as to support both backward and forward handover [25 – 27]. It enables the MT to instantaneously detach from its current AP and hand over its connection to the target AP at any time instance. As during backward handover the handover signaling is exchanged with the old AP, this instantaneous detach facility of the proposed scheme is of no particular use in this case. The disadvantage of these schemes is that it does not guarantee a lossless transfer of data during a soft handover. However, it can be very successfully exploited to provide zero cell loss forward handover [23].

Soft handover will be a basic feature for high-tier Mobile ATM systems; as air interface technologies for next generation wideband mobile services will enable soft handover, like for example 3rd generation CDMA technology [23].

Soft handover poses the highest requirements on the radio technology, because it requires the MT to be able to communicate concurrently with two APs. Therefore, in the overlapping boundary region, it enables dynamic selection of the best radio path. Provided that the overlapping region is sufficiently large and both APs can maintain a sufficiently strong signal in this region, this ensures enhanced QoS for the connection as well as handover reliability. Soft handover provides the best QoS performance of all handover alternatives [23]. The old and the new path are synchronized at the ATM level, facilitating path selection and therefore handover connection rerouting in real-time without introducing any delay or disruption into the cell stream. It is therefore the most attractive handover alternative for real-time services also requiring an outstanding cell loss performance during handover [31].

Two basic features must be taken into account during a soft handover process: synchronization of the two communication paths over the different APs and dynamic path selection. Both these functions should be realized with an efficient handover protocol.

2.7 Handover QoS requirements

The support of multimedia connections is one of the main objectives of WATM. Some multimedia connections require mainly low delay and delay variation while others require very low cell loss. For data applications, the most critical QoS factor is cell loss and is the most negative factor that influences the overall performance of the system. For WATM systems employing very small radio cells, and hence more APs, the handover rate could be very high. If even one ATM cell is lost during handover, the overall effect on the system throughput could also affect the delay experienced by the MT, if error correction mechanisms are used at the end systems. Therefore, one of the key design objectives for a handover in a WATM system is that it is lossless [23].

Interactive processes must also be catered for in a WATM system. For these services, the overall system delay becomes the most critical design objective. In order to ensure a lossless handover, one would have to introduce some kind of data buffering in the system during handover. The extra buffering would result in an increase in the system delays and needs to be minimized. The design objective for handover is thus to ensure a lossless handover while keeping the overall system delay as small as possible.

2.8 Cell routing and loss in the MES

As mentioned in chapter 1, the MES acts as the COS for Intra-switch handoff, and is located in the WATM network. Hence, the COS is a special case of the MES. In this section we will concentrate on the MES ability to rerouted connections internally as the COS does. By the saying that the MES acts as the COS, it is simply meant that the MES is the anchor in the handoff rerouting process.

The COS is an ATM switch which acts as an anchor in the handoff rerouting process. Most papers in the literature on handover in WATM networks concentrate on the design of handover signaling architectures without providing a design for the MES or COS [1, 6-8, 11-13, 25 -27]. To the author's knowledge, only four papers explicitly address the hardware design of the MES [15, 18, 29, and 30]. However, all these designs require a hardware change to be made to a regular ATM switch. The MES is designed such that cell buffering is done in the switch fabric in order to facilitate a lossless handoff. Therefore, a new buffer for handoff purposes is added to the switch hardware in order to avoid cell loss within the switch. Moving cells from one buffer to another increases the end-to-end delay and wastes memory bandwidth. To realize these functions with a regular ATM switch is not straightforward.

We view the WATM and fixed ATM networks as integral and fixed ATM switches are enhanced with mobility specific functions. The MES can support both fixed and wireless services, and no special types of switches are required to support mobility. The functionality required to support mobility is added to the fixed ATM switch as an external addition. The conversion from ATM to WATM cell is performed at the AP. Thus the MES is in essence only responsible for the rerouting of the connections during a handoff. We have stressed in chapter 1 that the MES should be 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 infrastructures. The solution is to make use of external buffers at the AP and MT [6 - 8, 11 – 13, 25 -27]. Thus, instead of performing cell buffering within the ATM switch thereby requiring a hardware change to the fixed ATM switch, the buffering is done elsewhere in order to simplify the process and to ensure the deployment of existing fixed ATM switches without any hardware changes.

With the 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.

For uplink data from the MT to the fixed ATM network, buffering is performed at the MT until a new connection is established during a handoff [6 - 8, 11 -13, 25 - 27].

For downlink traffic, buffering is required at the new AP until the new connection sub-path between the COS and a new AP is established by a handoff procedure. A handoff is usually initiated by a MT when it detects that the signal strength to the old AP is becoming too weak. An important step in a handoff procedure is to determine a COS.

and the new AP. As soon as the COS receives confirmation that the new sub-path is ready, it should stop transmitting ATM cells destined for a MT on the old path and send data via the new path. The MT will then be informed by the COS that it will no longer receive downlink data via the old path and receive a last packet indication message from the COS. This happens while the MT is still connected to the old AP (soft handover); hence downlink data is sent via the new path and has to be buffered at the new AP. Once the MT acquires a channel to the new AP, the buffered ATM cells are forwarded to the MT. The sequence of the handover signaling architecture should be precise in order to guarantee in-sequence cell delivery.

2.9 MES Requirements

We have argued and stressed that the design of the MES and handover signaling protocol should require no changes to switch hardware. Therefore, in order to accommodate handover in ATM switches, one needs to modify an external entity to the switch to support mobility functions [5]. Before we can address the design of the MES, we need to examine its requirements in more detail.

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

- The use of a handover signaling protocol should be supported by the MES. The signaling 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 signaling 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 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.

The messages that the MES receives is done via the handoff signaling protocol that is responsible for the sequence of events during a handoff. The MES should be designed to support a handoff protocol in order to facilitate a fast and lossless handoff.

2.10 Handover Performance Metrics

At this stage, it should be clear that the handover protocol must maintain lossless end-to-end connectivity during the handover process. The effectiveness of the rerouting performed by the handoff protocol is determined by several criteria [6, 8]. In particular, it is desirable to minimize the service disruptions (such as loss of motion in a video sequence due to the replay of the last frame when subsequent frames do not arrive on time) and overheads such as latency, MT and AP buffering, and excess reservation of network resources.

We can measure the performance of handover protocols using the following metrics:

- The service disruption time, the time during which the MT cannot transmit data
- Handoff completion time, the total time for all the handoff signaling to complete
- Buffering required at the AP for buffering of downlink data during a handoff
- Buffering required at the MT for buffering of uplink data during a handoff

The values of the metrics, evaluated for a given set of technology-dependent parameters and various parameters of network connections, will give some idea of the effectiveness of a handover protocol [6, 8]. $T_{disrupt}$ corresponds to the service disruption time, that is, the time during which the MT cannot receive data on its downlink. (This disruption may manifest itself as a pause in the playback of video.) BUF_{MT} and $BUF_{AP,down}$ correspond to the buffering required on the MT and the AP to prevent data from becoming out of order as a result of a handoff operation. $T_{complete}$, is the total time it takes for all the handover signaling to complete.

TABLE 2.2
TECHNOLOGY DEPENDENT NETWORK PARAMETERS

Symbol	Definition
BW_{wl}	Bandwidth of the wireless link
BW_{w}	Bandwidth of the wired backbone ATM network
L_{lw}	Latency of the wireless link
L_{w}	Latency of the wired backbone ATM network
PPT_{fixed}	Protocol processing time for signaling messages
PPT_{adm}	Protocol processing time where admission control tests needs to be performed
S_{sig}	Upper bound on the size of a signaling message
S_{data}	Maximum size of a data packet
$T_{acquire}$	Time for the MT to acquire a wireless channel to the AP

2.11 Analytical Analysis of the handover schemes

By making use of an analytical model of the handover protocol and a characterization of the WATM network, one can derive mathematical expressions for the metrics listed in the previous section [6, 8]. The analytical analysis uses a set of technology-dependent parameters, shown in Table 2.2, to characterize a fixed ATM and WATM network.

Each connection to and from a MT at the time of handoff can be characterized by the set of parameters listed in Table 2.3. These parameters are dependent on the locations of the MT as well as the topology of the network. By varying these parameters, we can examine the overheads associated with each handover protocol over varying lengths of connections.

TABLE 2.3
PARAMETERS CHARACTERIZING NETWORK CONNECTIONS

Symbol	Definition
H_{new}	Number of hops between the COS and the new AP
H_{old}	Number of hops between the COS and the old AP
H_{sig}	Number of hops between the old and new APs along the signaling channel

The analytical analysis provided in the literature assumes that $H_{new} = H_{old} = \frac{H_{sig}}{2}$ in order to simplify the analysis [6, 8]. However, the analysis fails to take advantage of the fact that the network is linear. For a linear network, if the time required to perform an operation for one hop is x seconds, then for two hops it will be 2x seconds, for three hops it will be 3x seconds, etc. Hence, by a linear network we mean that the time required to perform an operation with an increase in the number of hops will follow a straight line. Therefore, we provide an analytical analysis in this thesis that stipulates that H_{new} and H_{old} does not have to be equal and evaluate the performance metrics of section 2.7 without any such restrictions.

2.12 Summary of handover issues

One of the functions required 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.

Much of the research published on WATM have addressed the issues mentioned in this chapter, but to the authors knowledge, no effective method has addressed handover in such a manner that it be lossless and at the same time be part of a standard fixed ATM switch.

In the design of the MES two approaches can be adopted. The MES can be designed such that a hardware change is done to the switch fabric [15], or the MES can be designed such that any mobility functionality is implemented in an external processor. Since we have stressed that the MES should be efficient and yet simple enough to be implemented as a limited enhancement to a fixed ATM switch, the later approach will be used in the design.

A number of handover protocols have been designed to address the issue of handover in WATM networks [6 - 8, 11 – 13, 25 -27]. The handover protocols in the literature simply provide a signaling sequence for handover without specifying how the MES or the other WATM functional entities could be physically realized.

Chapter 3

Handover Signaling Analysis

3.1 Introduction

We mentioned in Chapter 1 that the focus of this thesis is on the design and implementation of the WATM functional architecture in order to facilitate a seamless handoff. Until now, the signaling sequence required by the handover rerouting schemes to facilitate a handoff, has not been investigated. Hence, this chapter describes the procedures for ATM-layer connection handover. The WATM control plane entities described in the previous chapter is responsible for facilitating the handover protocols where necessary.

The handover rerouting schemes to be used in WATM networks were standardised by the ATM forum wireless workgroup in BTD-WATM-01.13 [1]. The handover signaling protocols are based on the PR and PE schemes. However, the specification requires that the two schemes be used as a combination to facilitate the Two Phase Handoff scheme.

The mobile ATM specification [1] deals with ATM network support for connection handover. The essential approach used is to specify enhancements to existing ATM-layer functionality to support terminal mobility.

The soft handover model is considered and described in this chapter. Under this model, the M-UNI layer at the MT initiates handover for all active connections when it receives the "Handover Required" signal across *Inf.*1, Fig 2.1. The handover protocols presented are relevant for Inter-switch handoff. However, the protocols can also be adapted for Intra-switch handoffs. From the MT's perspective, this distinction is neither visible nor relevant. The procedures in the network, however, are different, i.e. less signaling messages are required to facilitate an Intra-switch handoff.

In the sections that follow, we present the signaling sequence of the PE and PR schemes together with an analysis. The analysis is performed by considering the signaling messages required to be exchanged during handoff in a common WATM network structure. A quantitative comparison of the PE and PR schemes are presented using analytically derived formulas for the performance metrics such as the: (i) Service disruption time, (ii) Handoff completion time, (iii) Buffer requirements at MT, and (iv) Buffer requirements at AP required during a handoff. These are important performance metrics to be considered as they affect the quality of service of the connections. The performance metrics will also be obtained from the evaluation platform and compared with the analytical analysis in chapter 6.

With the comparison of the PR and PE schemes we show how the two schemes can be combined to form the Two Phase handoff scheme.

3.2 Network Parameters used in the analysis

The parameters used in the analysis are given in Table 3.1. These parameters are only valid in the context of this chapter and are used to evaluate the performance metrics listed in the previous section.

It is assumed that the maximum throughput for a connection is limited by the throughput of the wireless connection. The analysis also assumes perfect delivery of signaling messages. We compute and measure the time taken for each of the signaling messages to be transmitted, forwarded, and processed where necessary. The calculated time represents average time scales. Based on the time taken by these signaling messages, the performance metrics will be obtained. Some of the parameters used are hardware and network specific (Table 3.1) [6, 8]. The analysis in this chapter is similar to the analysis in [6] and [8]. However, [6] and [8] assumes that $H_{new} = H_{old} = \frac{H_{sig}}{2}$ in order to simplify the analysis. The analysis fails to take advantage of the fact that the network is linear, this concept was explained in section 2.11. Therefore, we provide an analytical analysis in this chapter that stipulates that H_{new} and H_{old} does not have to be equal and evaluate the performance metrics of section 2.7 without any such restrictions.

TABLE 3.1
NETWORK PARAMETERS FOR HANDOVER SCHEMES

Parameter	Symbol	Value
Number of hops from the		
i) New AP to the COS	H_{new}	variable
ii) Old AP to the COS	H_{old}	variable
iii) old MES to new MES through Signaling Channel	H_{sig}	variable
Bandwidth of the Signaling Virtual Channel (SVC)	BW_{sig}	64kb/s
Bandwidth of the user data channel	BW_{data}	2Mb/s
Latency of the wireless link	L_{wl}	3ms
Latency of the wired backbone	L_{w}	500μs
Protocol processing time for signaling messages	PT_{fix}	0.3ms
Protocol processing time where COS discovery algorithm has to be performed at each node, in excess of fixed protocol processing time	PT_{cos}	1ms
Time for a MT to acquire a wireless channel to a base station	T_{acq}	2ms
Upper bound on the size of a signaling message (bytes)	S_{sig}	50
Max size of a data packet (bytes)	S_{data}	53
Protocol processing time where admission control is to be performed, in excess of fixed protocol processing time.	PT_{adm}	2ms

3.3 Signaling sequence and analysis of the PE Scheme

The handover signaling protocol presented in this section is based on the soft and backward handover classes for the Inter-switch PE scheme and is illustrated in Fig. 3.1. Soft handover requires the MT to be able to communicate concurrently with two APs. Therefore, in the overlapping boundary region, it enables dynamic selection of the best radio path. Provided that the overlapping region is sufficiently large and both APs can maintain a sufficiently strong signal in this region, this ensures guaranteed QoS for the connection as well as handover reliability. As described in chapter 2, soft handover provides the best QoS performance of all handover alternatives [23]. The old and the new path are synchronized at the ATM level, facilitating path selection and therefore handover connection rerouting in real-time without introducing any delay or disruption into the cell stream. It is therefore the most attractive handover alternative for real-time services [31].

The advantage of having a backward handover mechanism is that it gives one the possibility to choose the best AP for the MT to connect to.

In order to keep the explanation of the handover protocol simple, only one active connection to the MT is assumed and it can be either loss or delay sensitive. Where a distinction between delay and loss sensitive traffic is relevant, it will be explained in the context of the handover protocol.

Bit error rates in ATM networks can be expected to be in the order of 10^{-10} and in wireless networks in the order of 10^{-3} to 10^{-6}. We therefore do not expect any loss of signaling messages to occur in the fixed ATM network, however, this is not the case for a wireless network. If there is a loss or corruption of signaling messages in the wireless network, the MT and AP has to provide appropriate error correction mechanisms to ensure that the message is corrected or retransmitted. This will increase signaling delays.

The error control mechanisms used can be thought of as belonging to a sub-layer of the MAC layer (usually the upper part), referred to as Wireless Data Link Control (WDLC) sub-layer [55]. WDLC is responsible for recovering from occasional quality degradations of the wireless channel, and for providing an interface to the ATM layer in terms of frame format and required QoS. Error control techniques, in general, can be divided in two main categories, namely Automatic Repeat re-Quest (ARQ) and Forward Error Correction (FEC). In ARQ techniques, the receiver detects the erroneously received data and requests retransmission from the transmitter. Since retransmissions imply increased delays, ARQ is efficient for non-real-time data [56]. ARQ techniques are conceptually simple and provide high system reliability at the expense of some extra delay and bandwidth consumption due to retransmissions. FEC, on the other hand, is efficient for real-time data [57]. A number of bits are added in every transmitted data unit, using a predetermined error-correction code, which allows the receiver to detect and correct errors up to a predetermined number per data unit, without requesting any additional information from the transmitter. FEC techniques are fast at the expense of lower bandwidth utilization due to the transmission of additional bits. In wireless ATM, where both real-time and non-real-time data must be supported, a hybrid scheme combining

ARQ and FEC is usually used [58]. According to this, for real-time connections (e.g., CBR, rt-VBR) FEC bits are included in the header of every MAC data unit, to allow the receiver (AP or MT) to correct most of the errors. For non-real-time connections (e.g., nrt-VBR), no extra bits are included, and the AP (MT in the uplink) requests from the MT (AP in the downlink) the retransmission of erroneously transmitted MAC data units. However, error correction mechanisms for signaling are not in the scope of this project and we therefore assume perfect delivery of signaling messages.

The signaling sequence for the PE scheme is illustrated in Fig. 3.1.

The PVC between adjacent MES can be either a direct physical connection between two MES or it can be setup via the fixed ATM network. In either case, the principle of direct PVC between the MES is the same. Fig. 3.1 illustrates the handoff PVCs setup via the fixed ATM network.

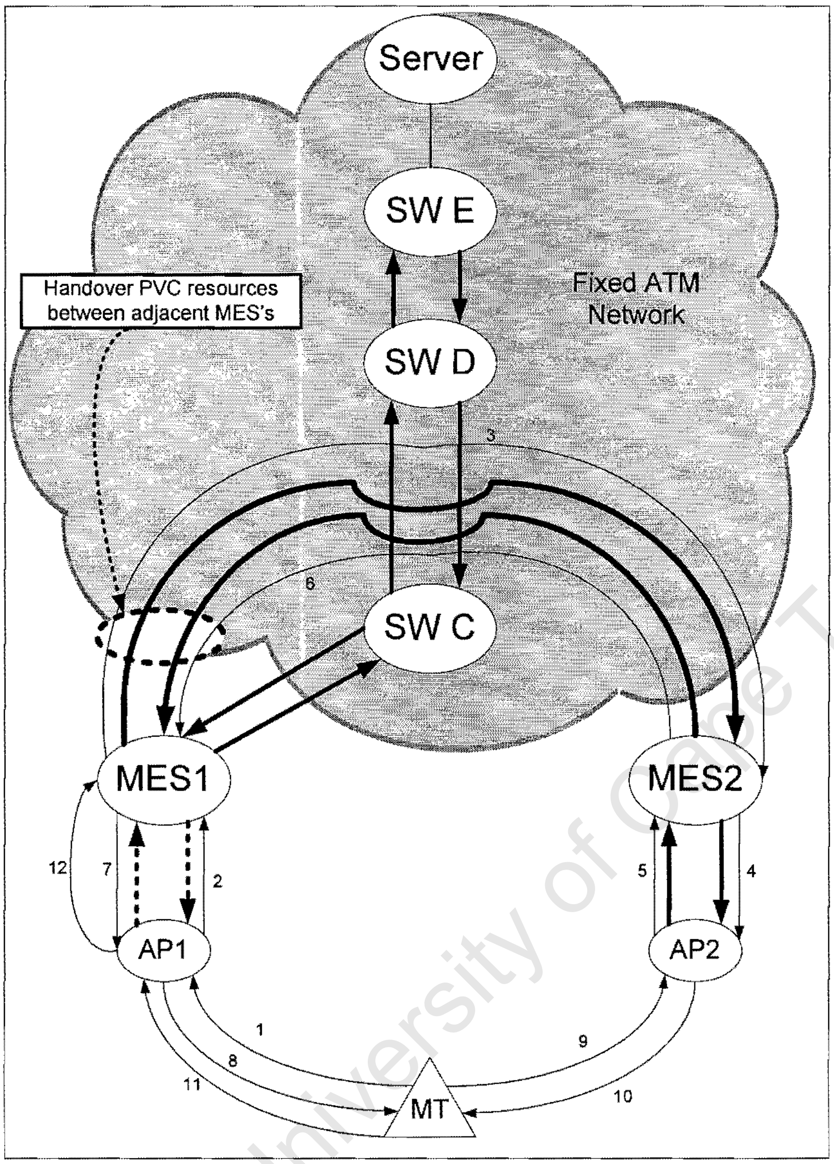


Fig. 3.1 Logical diagram of the PE Signaling sequence

The signaling sequence of the PE scheme is given below. In the text that follow, the numbers in parenthesis corresponds to the numbers of the signaling messages in Fig. 3.1.

1) The MT detects the handoff occurrence in advance by the radio receive signal dropping below a set threshold value and sends a handoff request (HO_REQUEST (1)) to the current serving AP, AP1. The time needed for this message, T_1, is the sum of the transmission time of a signaling message, the propagation time on the wireless link, and the fixed protocol processing time on AP1.

$$T_1 = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \quad (3.1)$$

2) After this the old AP sends the HO_REQUEST (2) to the old MES informing it that a handoff is requested by the MT. The message includes the identity of the new AP. The time needed for this message, T_2, is equal to the transmission time of the signaling message, the propagation time of the message on the fixed network for one link, and the fixed protocol processing time at MES1.

$$T_2 = \left(\frac{S_{sig}}{BW_{sig}} + L_w \right) + PT_{fix} \quad (3.2)$$

3) The old MES determines that the new AP is connected to an adjacent MES. The old MES checks if a PVC is available in its handover PVC resources, and if yes, sends a message (3) to the new MES requesting a VC (VC_REQUEST). The time needed for this message, T_3, is equal to the transmission time of the signaling message, the propagation time of the message on the fixed network for one link, the fixed protocol processing time at MES2, and admission control test that needs to be performed on the new request, which incurs a delay of PT_{adm}.

$$T_3 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} + PT_{adm} \right) \quad (3.3)$$

4) After the new MES has received message (3), it allocates local resources for the handoff and sends a RR_REQUEST message (4) requesting the new AP to allocate radio resources according to the expected QoS and bandwidth requirements. The time needed for this message, T_4, is equal to the transmission time of the signaling message, the propagation time of the message on the fixed network for one link, and the fixed protocol processing time at AP2.

$$T_4 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) \tag{3.4}$$

5) If resources are available at the new AP it will send a RR_CONFIRM message (5) as an acknowledgement of the SVC channel establishment from the new AP to the new MES. This signaling message retraces the path of message (4). Denote this time delay by T_5.

$$T_5 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) \tag{3.5}$$

The new AP will then wait for new downlink data and buffer them until the MT acquires a new radio channel to it.

6) Then the new MES sends a VC_CONFIRM message (6) to the old MES informing that the VCs have been allocated and that radio resources are available at the new AP. This signaling message retraces the path of message (3). Denote this time delay by T_6.

$$T_6 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) \tag{3.6}$$

7) The handoff can proceed, so the old MES informs the MT that handoff will be made to AP2 by sending the HO_RESPONSE message (7) to the old AP, and inserts a last packet indication (LPI) in the downlink data. This signaling message retraces the path of message (2). Denote this time delay by T_7.

$$T_7 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) \tag{3.7}$$

At the same time, the old down link direction of the connection is re-routed and extended by the old MES to the new AP2, through the new MES, using the pre-assigned PVC to bypass the traffic from PVC to SVC connections in the fixed ATM network.

8) Message (8) is the HO_RESPONSE message in the wireless link from the old AP to the MT.

$$T_8 = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.8}$$

The concept of "Tail" signals or Last Packet Indication (LPI) Cells was introduced by [9] to switch user data from an old VC to a new VC. The LPI cells are sent in-band on the same VC as the user cells. The LPI cells are distinguishable from the normal user cells by its payload content. In order to prevent a cell out of sequence problem, the LPI is used as follows: After the establishment of a new segment, the old MES sends a LPI cell in the downlink data to signal the switching of new downlink data to the MT. Upon reception of the LPI cell, the MT is assured that no more downlink cells will be received via the old downlink path. The LPI is only recognized by the MT and is treated as normal user data by the AP. The LPI cell will be discarded by the MT upon reception.

The reason why a LPI cell is needed in the downlink data is that the HO_RESPONSE message will arrive at the MT before the last data packet arrives. This is because the signaling VCs have priority in intermediate switches and will always be transmitted before user data. Although the HO_RESPONSE message indicates to the MT that the handover can proceed to the new AP, the MT can choose not to do so immediately. A reason for this is that the MT may elect to wait for the LPI to arrive before requesting a radio channel from the new AP thereby draining all downlink cells from the old path. This is important for data applications where the loss of even one cell could be detrimental to the MT applications. However, the MT may also elect to proceed with the handover immediately after receiving the HO_CONFIRM message without waiting for the LPI to arrive. In this case the MT might run applications that are very sensitive to delay and delay variation, and cell loss is not a crucial factor, e.g. video and voice applications.

If we assume that the MT application is loss sensitive with no delay requirements then the MT will wait for the LPI thereby ensuring that all downlink cells have arrived. While the MT waits for the LPI, the downlink cells that were sent to the new AP are buffered at the new AP until the MT acquires a channel to it.

(9) After receiving the LPI for the downlink data, the MT inserts a LPI packet in the uplink data path and then buffers new uplink data. At the same time the MT acquires a wireless channel to the new AP by sending RC_ACQUIRE message (9) to the new AP. This time delay is denoted by T_9.

$$T_9 = T_{acq} + \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.9}$$

(10) The new AP responds with a positive reply, RC_CONFIRM (10), since radio resources have been reserved for the handover. The time delay is denoted by T_{10}.

$$T_{10} = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.10}$$

At this stage the new AP starts transmitting the buffered downlink data to the MT.

(11) Once the MT receives the RC_CONFIRM message, it starts transmitting the buffered uplink data via the new AP and transmits a HO_COMPLETE (11) message to the old MES. . Denote this time delay by T_{11}.

$$T_{11} = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.11}$$

12) Message (12) is the handoff completion and connection release message from the old AP to MES1. The time delay is denoted by T_{12}.

$$T_{12} = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) \tag{3.12}$$

We assume that the MT application is loss sensitive with no delay requirements, hence the old MES will wait for the uplink LPI thereby ensuring that all uplink cells have arrived. As soon as the old MES receives the LPI for the uplink data, it will reroute the uplink data VC via the new AP. This concludes the PE handover protocol. However, at this stage, the old MES can request a route optimization to the new MES. This forms the second phase of the Two Phase Handoff Scheme and is an adaptation of the Partial Re-establishment (PR) Scheme which will be discussed in the next section.

The difference between the handover protocol described above and that found in [34] is that we buffer the new downlink data at the new AP and it is done at the old MES in [34]. The principle of buffering new downlink data is the same for both cases; however, if buffering is done at the old MES, it has to be performed before the data reach the output port of the switch. Hence, a new buffer has to be added between the switch fabric of the ATM switch and its output port. This adds extra complexity to the MES design and requires a hardware addition to the ATM switch. This is against our initial design objective that specifies that there should be no hardware changes to the fixed ATM switch.

Based on the time taken by the above signaling messages, various performance measurements can be derived. Expressions for four measures are given. 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 signaling 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.

$$T_{disrupt} = T_9 + \frac{S_{sig}}{BW_{sig}} + \frac{S_{data}}{BW_{data}} + L_{wl} \tag{3.13}$$

The handoff completion time $T_{complete}$ 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 previous 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. Hence,

$$\begin{aligned} T_{complete} &= T_1 + \frac{S_{sig}}{BW_{sig}} + T_3 + T_4 + T_5 + T_6 + T_7 + T_8 \\ &+ \max\left[\left(\frac{S_{sig}}{BW_{sig}} + T_9 + T_{10}\right), \left(T_9 + T_{11} + T_{12} - L_{wl} + PT_{fix}\right)\right] \end{aligned} \tag{3.14}$$

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.

$$BUF_{MT} = (T_9 + T_{10})BW_{data} \tag{3.15}$$

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.

$$BUF_{AP,down} = \left(T_7 + T_8 + \frac{S_{sig}}{BW_{sig}} + T_9 + \frac{S_{sig}}{BW_{sig}}\right)BW_{data} \tag{3.16}$$

The time signaling sequence of the PE scheme is illustrated in Fig. 3.2. The diagram includes an illustration of the performance metrics listed above.

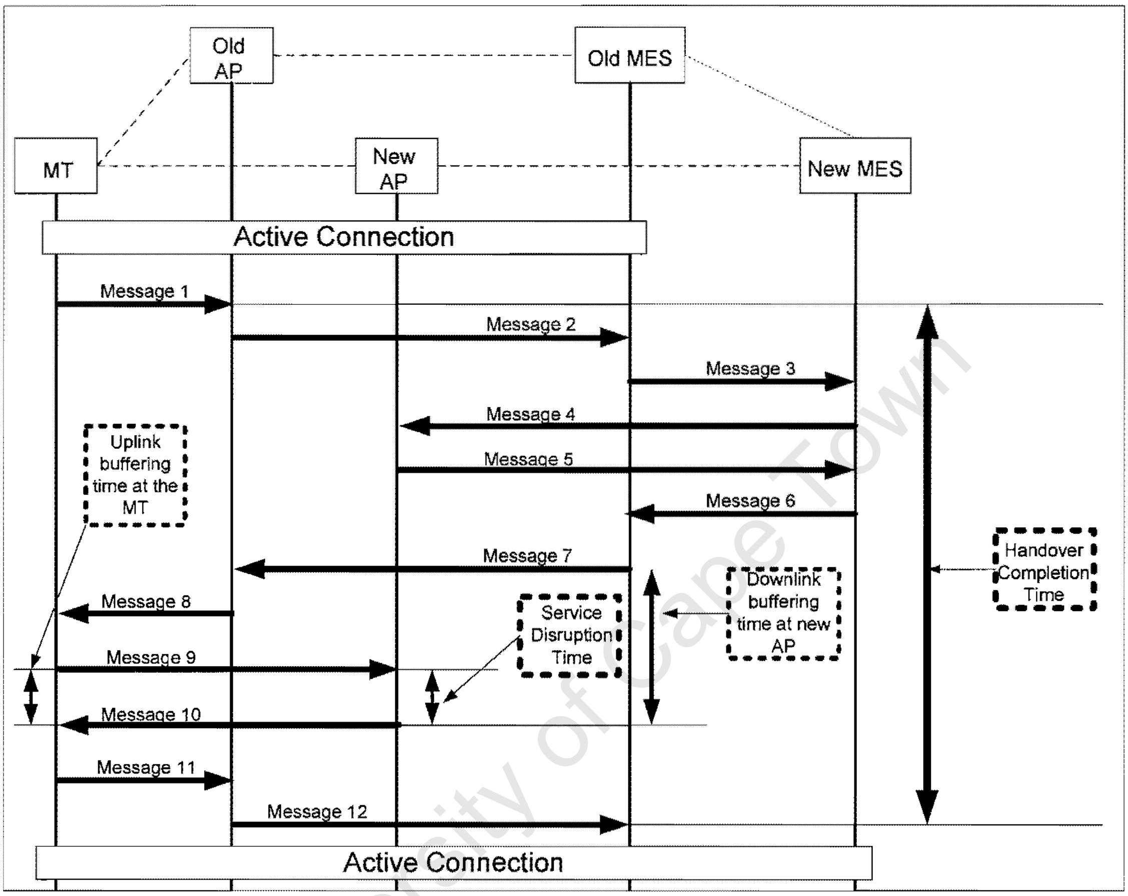


Fig. 3.2 PE signaling sequence with time perspective

3.4 Signaling sequence and analysis of the PR Scheme

Once again, in order to keep the explanation of the handover protocol simple, only one active connection to the MT is assumed and it can be either loss or delay sensitive. Where a distinction between delay and loss sensitive traffic is relevant, it will be explained in the context of the handover protocol. We also assume perfect delivery of signaling messages.

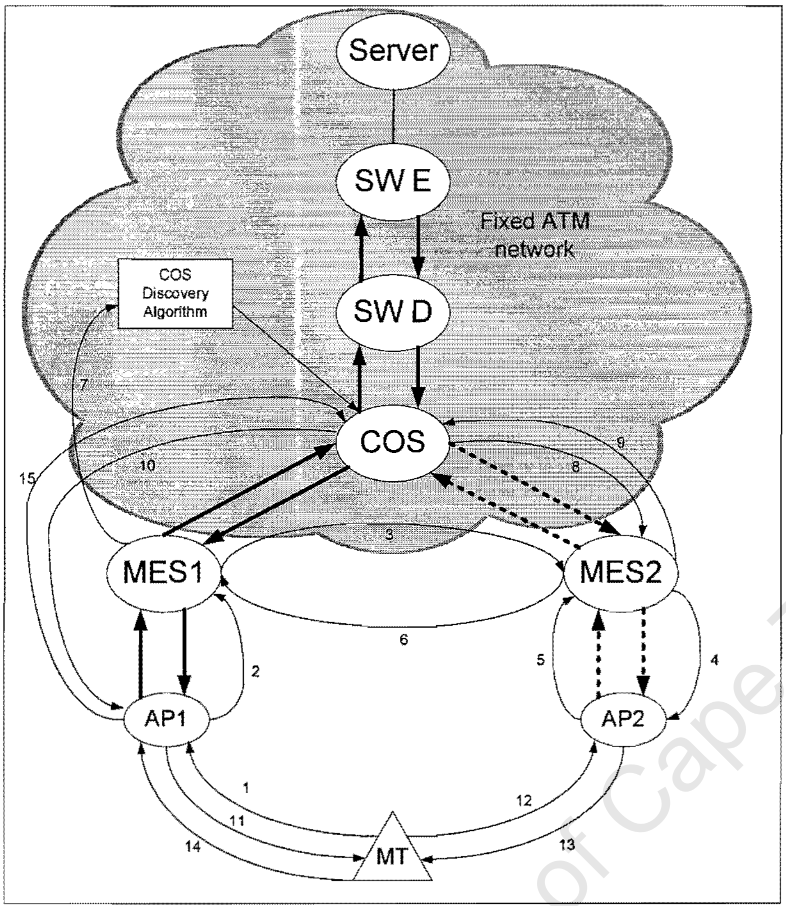


Fig. 3.3 Signaling sequence for the Partial Re-establishment Scheme.

(1) Consider Fig. 3.3 where the handover is initiated by the radio receive signal at the MT dropping below a set threshold value. To request a handover, the MT sends a handover request message (1) (HO_REQUEST) to the old AP, AP1. The HO_REQUEST message includes the identity of the new AP. The time needed for this message, T_1, is the sum of the transmission time of a signaling message, the propagation time on the wireless link, and the fixed protocol processing time on AP1.

$$T_1 = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.17}$$

2) After this the old AP sends the HO_REQUEST (2) to the old MES informing it that a handoff is requested by the MT. The message includes the identity of the new AP. Its total time is the end-to-end delay through the signaling VC between AP1 and MES1 plus the fixed protocol processing time at MES1. This time delay is denoted by T_2.

$$T_2 = \left(\frac{S_{sig}}{BW_{sig}} + L_w \right) + PT_{fix} \tag{3.18}$$

(3) After receiving the HO_REQUEST message from the old AP, the old MES contacts the new MES via a signaling channel between them to request local resources from the new MES and radio resources from the new AP. This is done with the VC_REQUEST message (3). The time taken for this is equal to the transmission and propagation time in the signaling channel between the old and new MES, plus the fix protocol processing time at the new MES.

$$T_3 = \left(\frac{S_{sig}}{BW_{sig}} + L_w \right) H_{sig} + PT_{fix} \tag{3.19}$$

(4) After the new MES has received message (3), it allocates local resources for the handoff and sends a RR_REQUEST message (4) requesting the new AP to allocate radio resources according to the expected QoS and bandwidth requirements. Its total time is the end-to-end delay through the signaling VC between AP2 and MES2 plus the fixed protocol processing time at AP2. This time delay is denoted by T_4.

$$T_4 = \left(\frac{S_{sig}}{BW_{sig}} + L_w \right) + PT_{fix} \tag{3.20}$$

5) If resources are available at the new AP it will send a RR_CONFIRM message (5) as an acknowledgement of the SVC channel establishment from the new AP to the new MES. If no resources are available at the new AP, a negative response will be issued and the call will be dropped. The MT will then have to request a new connection setup at the

new AP. For more information on call blocking and blocking probability, the reader is advised to consult [13]. We assume that resources are available. The time delay is equal to T_5. Hence,

$$T_5 = \left(\frac{S_{sig}}{BW_{sig}} + L_w \right) + PT_{fix} \tag{3.21}$$

The new AP will then wait for new downlink data and buffer them until the MT acquires a new radio channel to it.

6) The new MES sends a VC_CONFIRM message (6) to the old MES informing that the VCs have been allocated and that radio resources are available at the new AP. This message retraces the path of message 3. Thus,

$$T_6 = \left(\frac{S_{sig}}{BW_{sig}} + L_w \right) H_{sig} + PT_{fix} \tag{3.22}$$

(7) On reception of the VC_CONFIRM message from the new MES, the old MES invokes a COS discovery algorithm in the fixed ATM network in order to discover the most optimal path between itself and the new MES. The messages needed to find the COS will incur a time equal to the transmission and propagation time along each hop, plus the protocol processing time to perform the COS discovery algorithm in each switch along the path between the old MES and the COS. Denote this time delay by T_7.

$$T_7 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{\cos} \right) H_{old} \tag{3.23}$$

The COS algorithm process determines the trade-off between end-to-end routing efficiency and handoff latency. For an optimal end-to-end route, a complicated algorithm is needed for selection of a COS. This will increase handoff latency [20]. A list of COS positions can also exist between the old MES and the new MES for rapid use in case of a handover without the need for a COS discovery algorithm [35]. However, COS discovery algorithms is not in the scope of this project.

(8) After the COS is identified, a new path is established to the new MES by the COS sending VC_REQUESTS (8) to intermediate switches. Admission control tests are carried out at each switch along each hop, plus the fixed protocol processing time at each switch along the path. Denote this time delay by T_8.

$$T_8 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} + PT_{adm} \right) H_{new} \tag{3.24}$$

Handover will be granted if sufficient network resources exist at intermediate switches. We assume that sufficient resources exist at intermediate switches and that the radio resources at the new AP are the only scarce resource. However, the radio resources at the new AP were already granted. Thus the handover can proceed.

(9) Message (9) is the acknowledgement of the partial channel establishment from MES2 to the COS. This signaling message retraces the path of the partial establishment, hop by hop. Denote this time delay by T_9.

$$T_9 = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) H_{new} \tag{3.25}$$

After the COS receives the VC_CONFIRM message (9), it inserts a Last Packet Indication (LPI) in the payload of the last cell in the down link data and switches the newly arriving downlink data to the new AP.

At the same time, the COS transmits a Handover Confirm (HO_CONFIRM) (10) message to the old AP.

(10) The HO_CONFIRM message (10) indicates to the AP that the LPI has been inserted into the downlink data path and that the handover can proceed. This time delay is denoted by T_{10}.

$$T_{10} = \left(\frac{S_{sig}}{BW_{sig}} + L_w + PT_{fix} \right) H_{old} \tag{3.26}$$

The LPI packet indicates to the old AP and MT that no more downlink data will be arriving via the old path. The reason why a LPI is needed in the downlink data is that the HO_CONFIRM message will arrive at the MT before the last data packet arrives. This is because the signaling VCs have priority in intermediate switches and will always be transmitted before user data, as with the PE scheme. Although the HO_CONFIRM message indicates to the MT that the handover can proceed to the new AP, the MT can choose not to do so immediately. A reason for this is that the MT may elect to wait for the LPI to arrive before requesting a radio channel from the new AP thereby draining all downlink cells from the old path. This is important for data applications where the loss of even one cell could be detrimental to the MT applications. This satisfies our requirement specifying that there should be no loss of data for loss sensitive traffic. However, the MT may also elect to proceed with the handover immediately after receiving the HO_CONFIRM message. In this case the MT might run applications that are very sensitive to delay and delay variation, and cell loss is not a crucial factor, e.g. video and voice applications.

We assume again that the MT application is loss sensitive with no delay requirements. Thus, the MT waits for the LPI thereby ensuring that all downlink cells have arrived. While the MT waits for the LPI, the downlink cells that were sent to the new AP are buffered there until the MT acquires a channel to it. After receiving the LPI for the downlink data, the MT inserts a LPI packet in the uplink data path.

11) Message (11) is the HO_CONFIRM message in the wireless link from AP1 to the MT. The time delay is denoted by T_{11}.

$$T_{11} = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.27}$$

(12) The MT requests a radio channel from the new AP by sending a RC_REQUEST message (12). This time delay is denoted by T_{12}.

$$T_{12} = T_{acq} + \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.28}$$

This message also indicates to the new AP that the MT is ready to receive downlink cells. At this point, the MT inserts a LPI for the uplink data and starts to buffer new uplink data.

(13) The new AP responds by sending a RC_CONFIRM message (13) back to the MT. The time delay is denoted by T_{13}.

$$T_{13} = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.29}$$

After sending the RC_CONFIRM message to the MT, the new AP immediately starts to transmit the buffered downlink cells to the MT. Meanwhile the COS has received the LPI packet for the uplink data and switches the uplink data path via the new AP. As soon as the COS receives the LPI cell for the uplink path, it sends an additional message to the MT via the new AP indicating that the LPI has been received for the uplink path. The MT would then wait until it receives this message before it transmits the buffered uplink data via the new AP. We again specify that the MT may choose when to transmit the buffered uplink data. For delay sensitive applications, the MT will not wait for the LPI receive confirmation (LPI_CONFIRM) from the COS and will start to send the buffered uplink data to the new AP immediately after receiving the RC_CONFIRM message from the new AP. For loss sensitive data applications, the MT will wait for the LPI_CONFIRM message from the COS before transmitting the buffered uplink data to the new AP.

(14) Once the MT receives the RC_CONFIRM message from the new AP, it transmits a HO_COMPLETE message (14) to the old AP. The time delay is denoted by T_{12}.

$$T_{14} = \frac{S_{sig}}{BW_{sig}} + L_{wl} + PT_{fix} \tag{3.30}$$

(15) Message (15) is the handoff completion and connection release message from the old AP to the COS hop by hop. The time delay is denoted by T_{15}.

$$T_{15} = \left(\frac{S_{sig}}{BW_{sig}} + L_{w} + PT_{fix} \right) H_{old} \tag{3.31}$$

Based on the time taken by these signaling messages, expressions for the four performance metrics are given. 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 signaling 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.

$$T_{disrupt} = T_{12} + \frac{S_{sig}}{BW_{sig}} + \frac{S_{data}}{BW_{data}} + L_{wl} \tag{3.32}$$

The handoff completion time $T_{complete}$ 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. Hence,

$$T_{complete} = T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8 + T_9 + T_{10} + T_{11}$$

$$+ \max\left[\left(\frac{S_{sig}}{BW_{sig}} + T_{12} + T_{13}\right), \left(T_{14} + T_{15}\right)\right] \tag{3.33}$$

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.

$$BUF_{MT} = \left(T_{12} + T_{13}\right)BW_{data} \tag{3.34}$$

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.

$$BUF_{AP,down} = \left(T_{10}\right)BW_{data} + \left(T_{11} + T_{12} + T_{13}\right)BW_{data} \tag{3.35}$$

The signaling sequence for the PR scheme with a time perspective of the performance parameters is shown in Fig. 3.4. The X indicates the time at which the downlink and uplink data are switched.

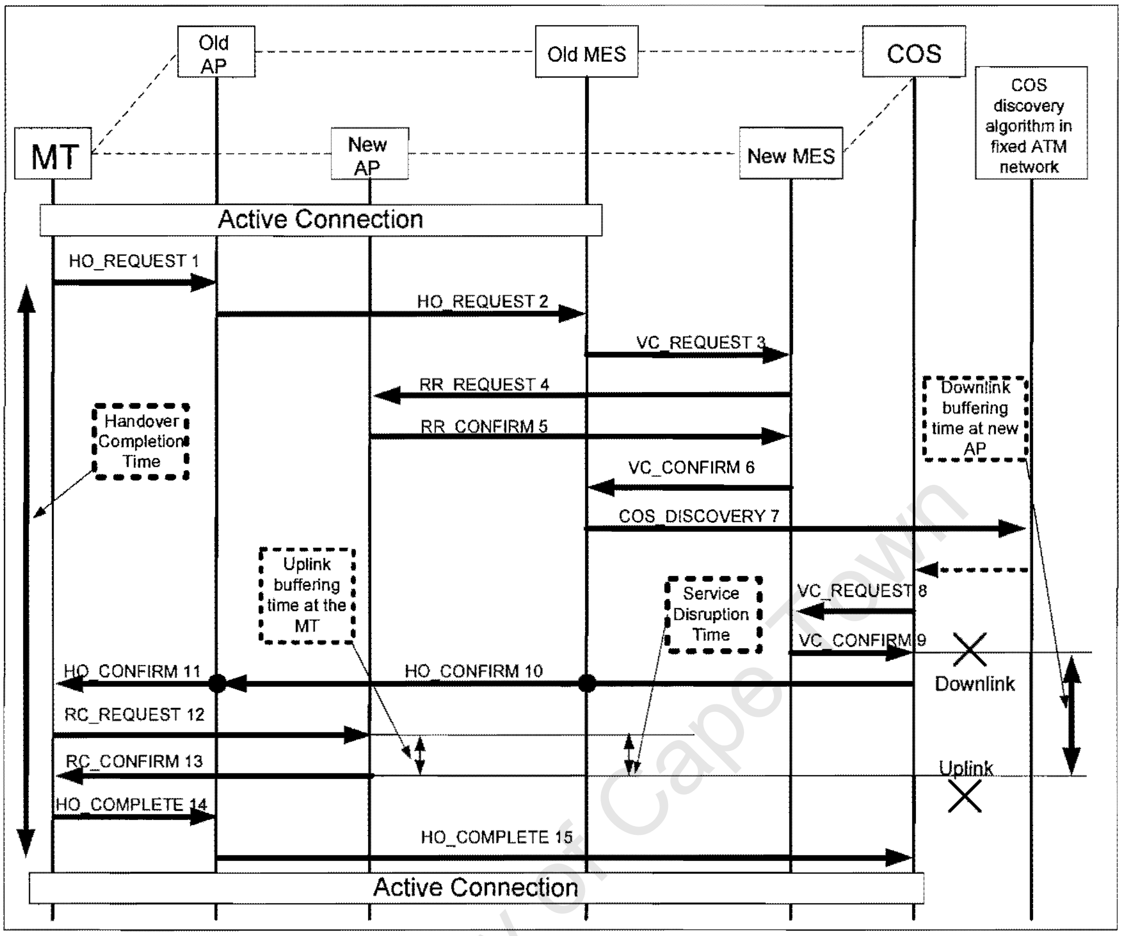


Fig. 3.4 PR signaling sequence with time perspective

3.5 Comparison of the analytical results

For evaluation of the analytical model it is initially assumed that $H_{new} = H_{old} = \frac{H_{sig}}{2}$ for the PR scheme, in order to simplify the calculations. However, H_{new} need not necessarily be equal to H_{old}. An alternative and more flexible analysis stipulates that H_{new} and H_{old} does not have to be equal and one could evaluate the performance metrics without any such restrictions.

The values for $T_{disrupt}$, $T_{complete}$, BUF_{MT}, BUF_{AP}, and R_{excess} depend on the specific hardware being used and also the number of hops being considered. Since the same hardware will be used to evaluate the various handoff schemes, comparisons can be made.

Values for the Service Disruption Time, Handoff Completion Time, Buffer Requirements at APs, and Buffer requirements at the MT were obtained by inserting the equations of sections 3.2 and 3.3 into Microsoft Excel with the network parameters specified in table 3.1. A discussion of the results is now presented.

Figure 3.5, 3.6, 3.7 and 3.8 show the performance metrics of the rerouting schemes as obtained from the analysis in graphical form.

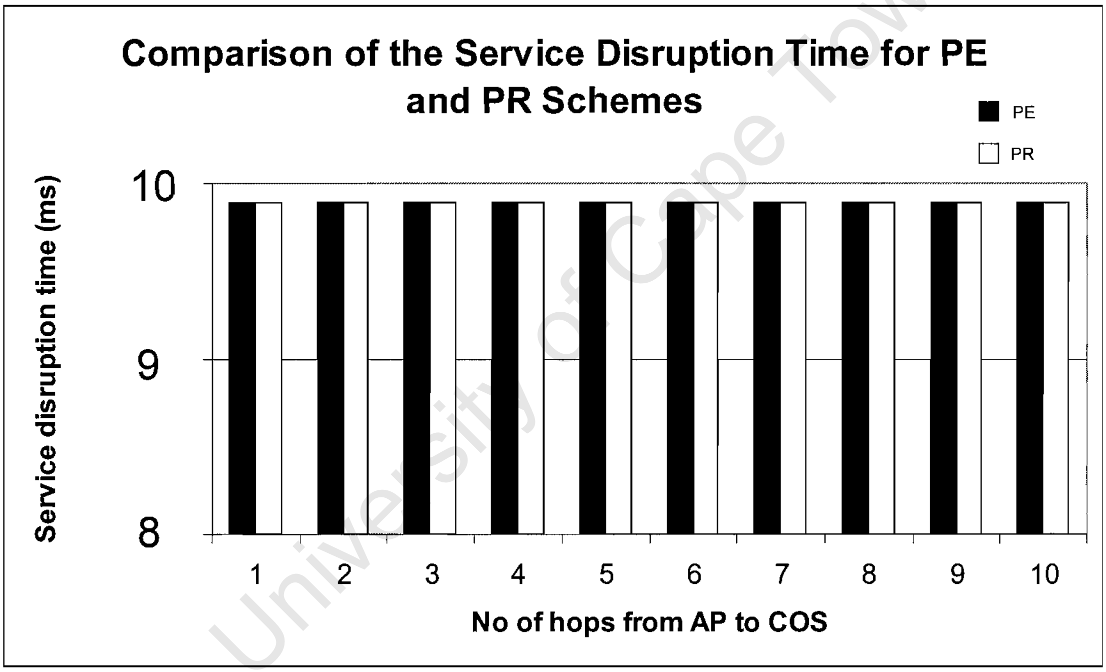


Fig. 3.5 Service disruption time for PR and PE Schemes

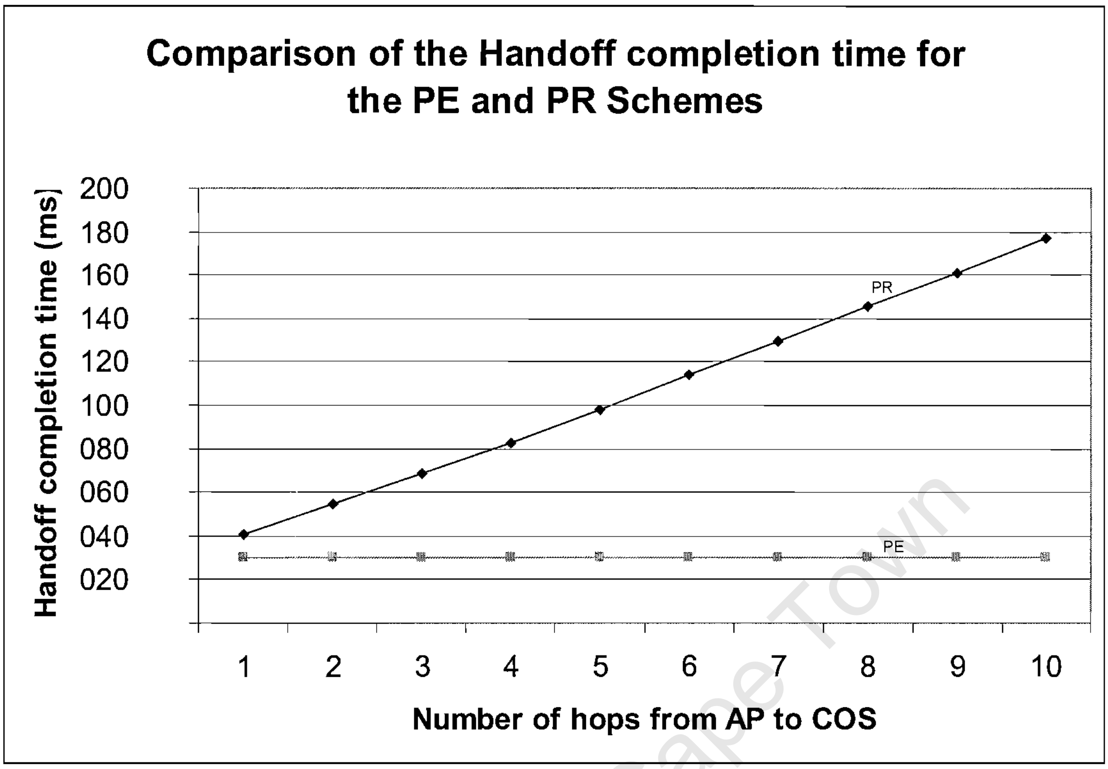


Fig. 3.6 Handoff completion time PR and PE Schemes

Fig. 3.5 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 signaling 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.

The handoff completion times of the Partial Re-Establishment and Path Extension Schemes are compared in Fig. 3.6. The Path Extension scheme is superior to the Partial Re-Establishment Scheme as far as the handover completion times are concerned. Hence, handover network signaling 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, Fig. 3.7. 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. 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.

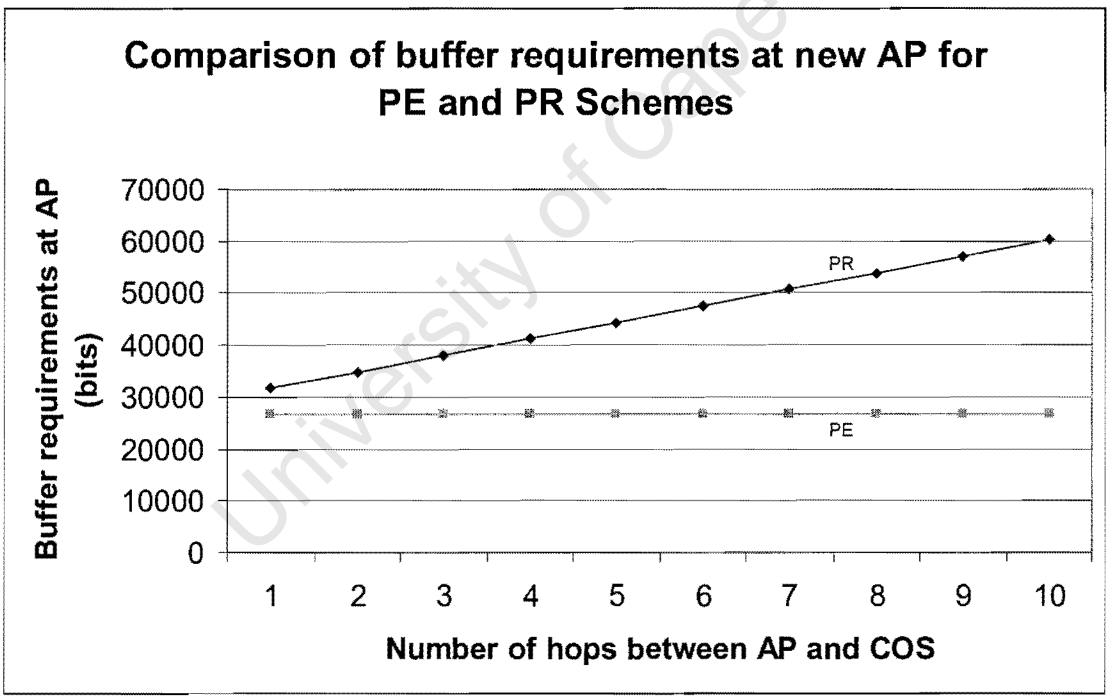


Fig. 3.7 Buffer requirements at AP for downlink during handover

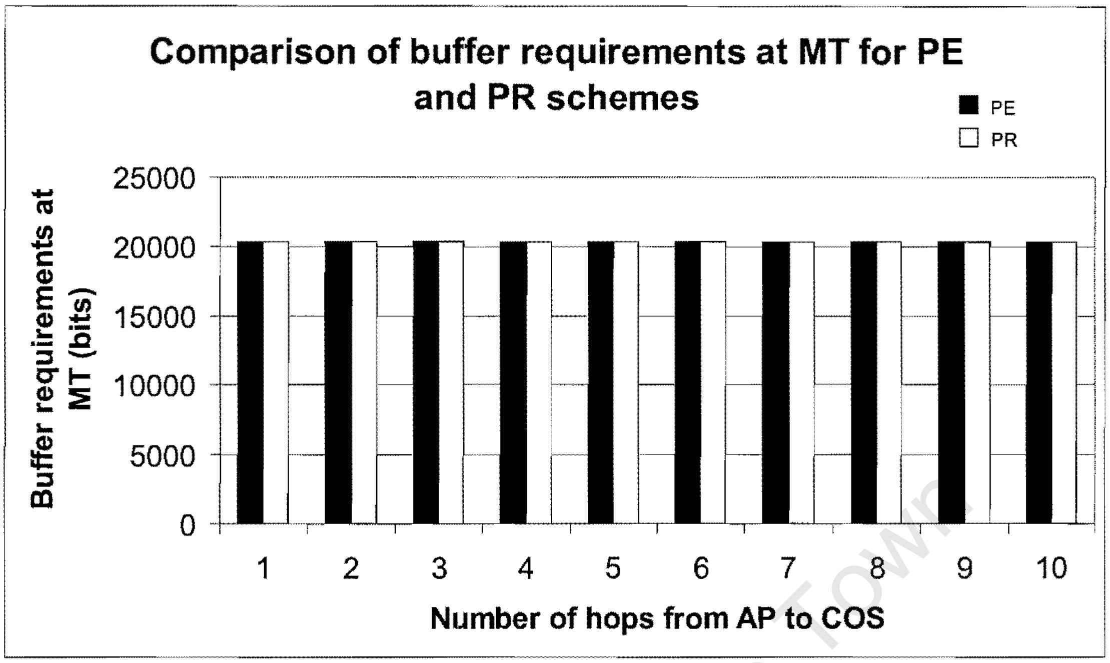


Fig. 3.8 Buffer requirements at the MT for the uplink during handover

The buffer requirements at MT for uplink data during handoff are the same for both schemes, Fig. 3.8. 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 9 and 10, Fig. 3.1 and messages 12 and 13, Fig. 3.3).

Consider Fig. 3.6, the new connection setup time is proportional to the number of switches between the COS and the AP. Therefore, as the number of hops between the COS and AP increase, so does the handover completion time. We have assumed that $H_{new} = H_{old} = \frac{H_{ctrl}}{2}$ for the PR scheme. We now take advantage of the fact that the network is linear. Consider the following scenario:

If we have x hops between the new AP and the old AP, with the COS y hops from the new AP and (x – y, x > y) hops from the old AP, the overall handover completion time will still be for x hops, since the network is linear. Hence, it does not matter where in the path the COS is located, the handover completion time is only dependent on the total number of hops between the new AP and the old AP. This scenario is illustrated in Fig. 3.9.

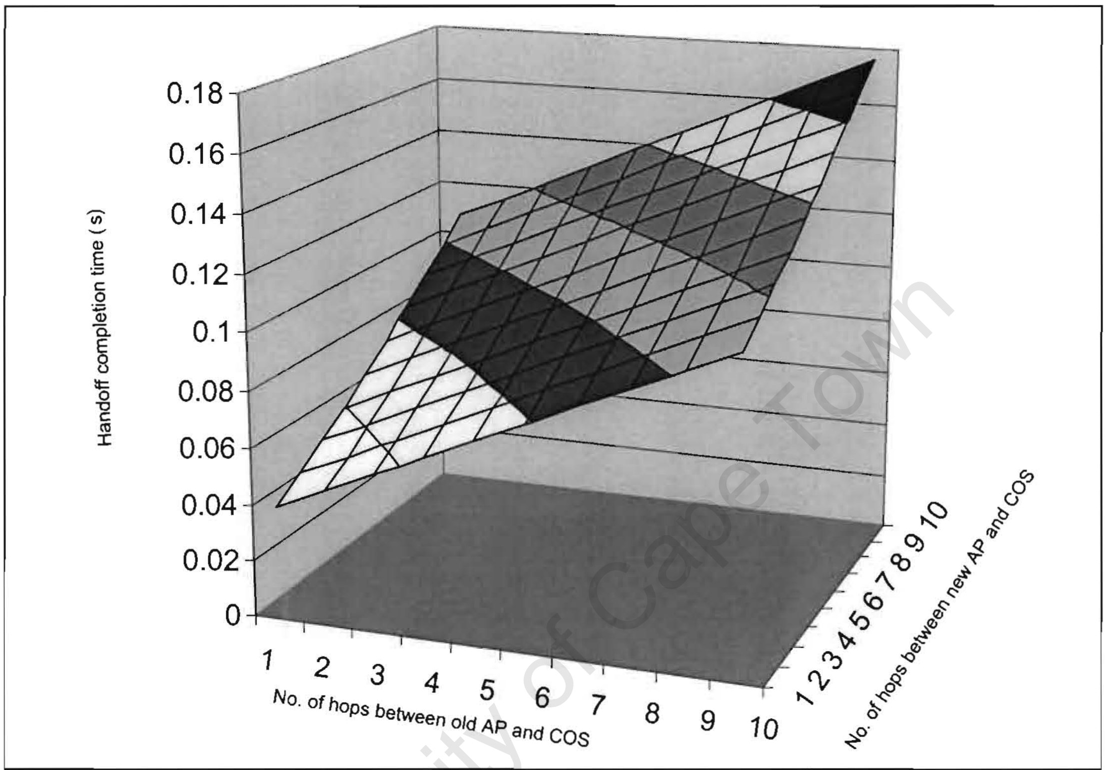


Fig. 3.9 Handoff completion time for the PR scheme with COS location varied

Fig. 3.10 shows the buffer requirements at the new AP for the downlink data by making use of the scenario described above.

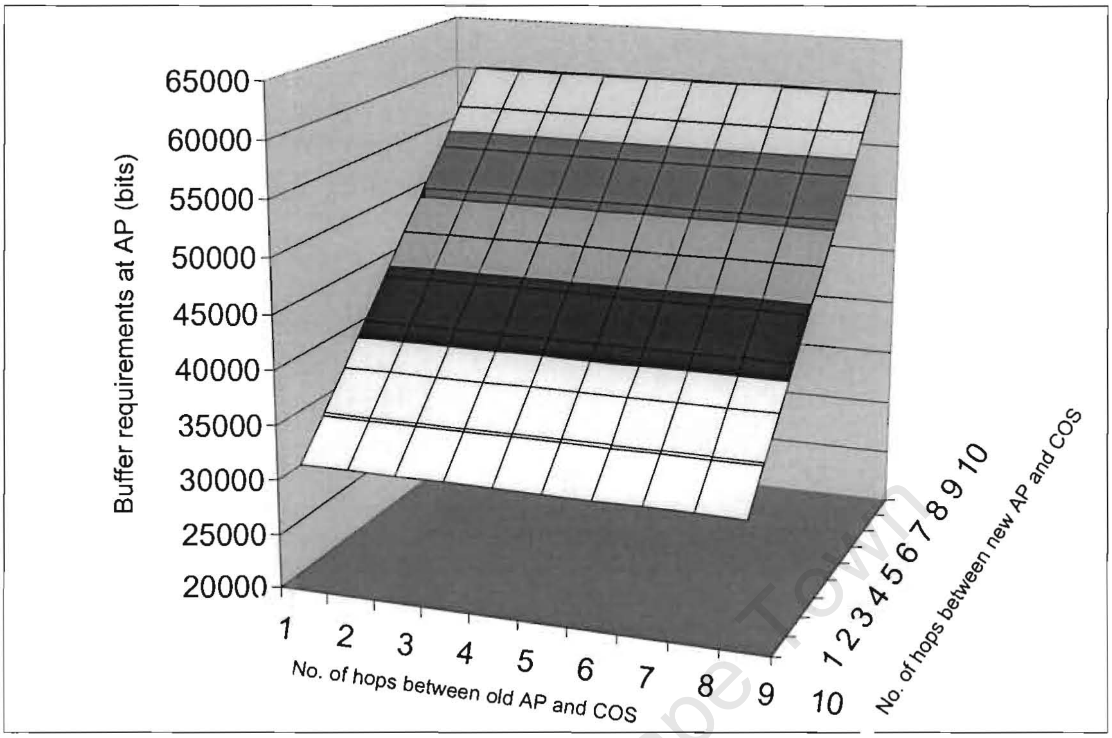


Fig. 3.10 Buffer requirements at the new AP with the COS location varied

The buffer requirements at the new AP for the downlink data is only dependent on the number of hops between the COS and the new AP and the number of hops between the old AP and the COS has no effect on it. By increasing the number of hops between the old AP and the COS, we only increase the handover completion time. By comparing Fig. 3.9 and 3.10, one can conclude that the performance of the PR scheme is enhanced when the position of the COS is closer to the new AP. Although we are not concerned with COS location algorithms, this result could be used in COS location algorithms in order to enhance the performance of the PR scheme.

3.6 The Two Phase Handoff Scheme

By comparing the characteristics and performance metrics of the PE and PR Schemes, one can conclude why the Two Phase Handoff Scheme was proposed in the literature. The Two Phase Handoff Scheme combines the PE and PR 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 compared to the PR scheme. However, for multiple switch handoffs, the PE scheme simply extends the data path from one MES to the next resulting in an inefficient path. 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.

The first phase of the two phase handoff scheme is exactly the same as the signaling sequence for the PE scheme presented in section 3.2. As soon as the PE signaling sequence has been completed, the old MES requests a route optimization in the fixed ATM network in order to locate the COS. This is part of the signaling sequence of the PR scheme presented in section 3.3. However, the second phase of the two phase handoff scheme starts at message number (7) of the PR scheme. This is because the radio resources at the new AP have already been reserved in the first phase of the two phase handoff scheme. In the signaling sequence of the PR scheme, the COS sends a LPI for the downlink data to the MT indicating that no more data will arrive and that the handover can proceed; after establishing a path to the new MES. However, in the second phase of the two phase handover protocol, the COS sends the LPI for the downlink to the new MES via the path created in the PE scheme (first phase). Hence, the MT is not aware of the optimization process.

As soon as the new path has been optimized, the PVC between the old and new MES is released.

Chapter 4

System design of the handover signaling entities

4.1 Introduction

The previous chapter presented and analysed the PE and PR handover signaling protocols, where several entities were identified to support the handoff. This chapter discusses the design and architecture of those entities. The most important design considerations were:

- Handover support should not be embedded in the ATM switch hardware.
- The use of a handover signaling protocol should be supported by the handover signaling entities.
- The system is required to provide in-sequence cell delivery of user data during a handoff.
- The MES should be able to switch ATM cells to a new VC in order to facilitate a handoff. The MES therefore requires direct access to the ATM switch routing tables in order to change it when requested by a handover operation.
- The new AP should buffer the downlink data until the MT acquires a channel to the new AP.
- The MT should buffer uplink data before acquiring a radio channel to the new AP.
- The design should support a variable number of hops between the MES and COS.
- Signaling messages should obtain higher priority for transmission at intermediate network nodes than user data.

The current design only supports one handover instance at any point in time. If more than one handover are requested simultaneously by multiple MTs connected to the same AP, each individual handover procedure would follow the same sequence of events. However, an additional module would have to be incorporated at the AP in order to provide a scheduling mechanism for simultaneous handoff requests. An AP scheduling architecture would have to support and guarantee the QoS of connections from various MTs. However, AP scheduling architectures is beyond the scope of this thesis.

4.2 WATM Functional User and Control Plane Entities

The WATM functional User and Control Plane architecture is defined in terms of various entities. This is shown in Fig. 4.1.

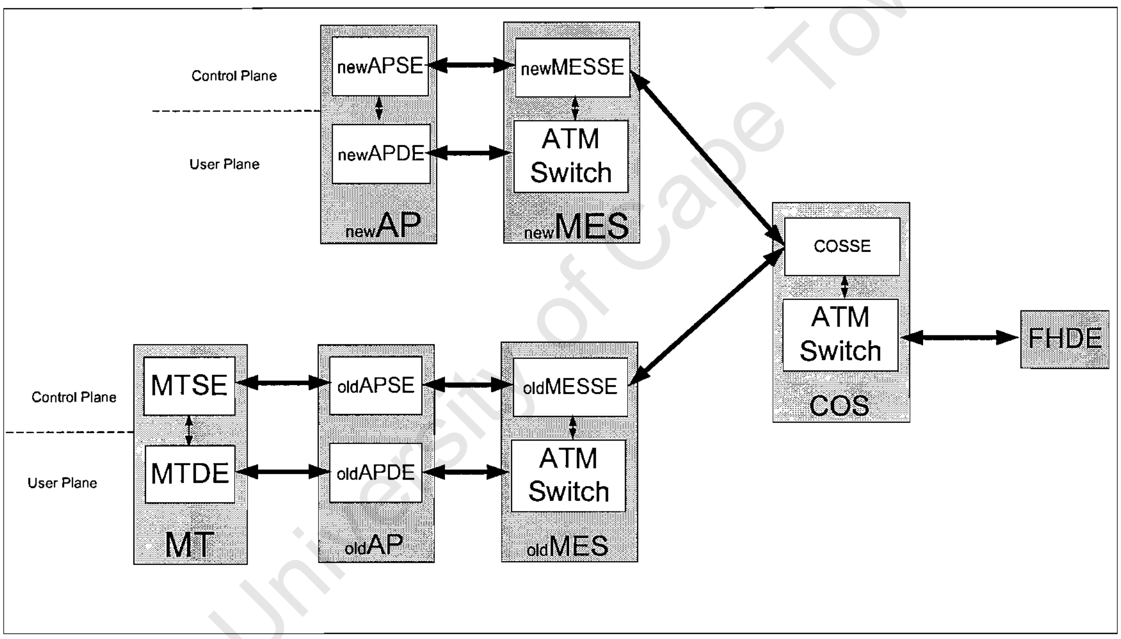


Fig. 4.1 Logical diagram of Control and User Plane entities

This section describes the functions of the various entities as illustrated in Fig. 4.1.

4.2.1 MTSE (MT Signaling Entity)

This functional entity is responsible for:

- Deciding when active connections must be handed over to a new network point of attachment (AP)

- Initiating handover for active connections
- Executing the handover protocols at the MT
- Interacts with the MTDE in order to request that user data be buffered and inserts a LPI message in-band with user data for loss sensitive traffic

4.2.2 MTDE (MT Data Entity)

This entity performs user plane functionality. It houses the user applications and provides an interface to lower layer transport protocols, i.e. ATM in this instance. The MTDE implements a buffer for uplink data when requested to do so by the MTSE.

4.2.3 APSE (AP Signaling Entity)

This functional entity performs the complimentary functions related to handover at the AP. The functions of the APSE at the old and new AP are different as far as the handover protocol procedures are concerned. The functions at the APSE include:

- Receiving handover requests from the MT and passing it on to the MES
- Executing the AP handover procedures
- Indicate to the AP Data Entity (APDE) when to buffer downlink data
- Indicating handover results to the MT
- Assign radio channels to the MT

4.2.4 APDE (AP Data Entity)

As with the MTDE, the APDE performs user plane functionality. It extracts the user data from the wireless medium and transmits it to the wired backbone network. The APDE at the new AP implements a buffer for downlink data when requested by the APSE.

4.2.5 MESSE (MES Signaling Entity)

This functional entity performs the complimentary functions related to handover at the MES. The functions of the MESSE at the old and the new MES are also different as far as the handover protocol procedures are concerned. These functions include:

- Receiving handover requests from the AP
- Deciding on which network point of attachment the connections must be handed over to
- Executing the network side handover procedures
- Indicating handover results to the AP
- Requesting radio resources from the AP

The MESSE has direct access to user data paths and is able to change the switch routing table to facilitate a handover. The MESSE is also able to generate and insert the LPI in-band in the user data paths.

4.2.6 COSSE (COS Signaling Entity)

This functional entity performs the complimentary functions related to handover at the COS. These functions include:

- Receiving handover requests from the MES;
- Deciding on which network point of attachment the connections must be handed over to;
- Executing the network side handover procedures; and
- Indicating handover results to the MES.

The COS is part of the fixed ATM network, hence, mobility signaling needs to be added to the fixed PNNI signaling protocol to facilitate the handover signaling in the fixed ATM network. Provision has been made in the WATM specification BTD-WATM-01.13 [1] for this protocol and it is referred to as M-PNNI.

4.3 WATM Protocol Layering

Fig. 4.2 depicts the protocol layers required to support the logical configuration shown in Figure 4.1. The functions of the layers are identical to those described in section 2.4.

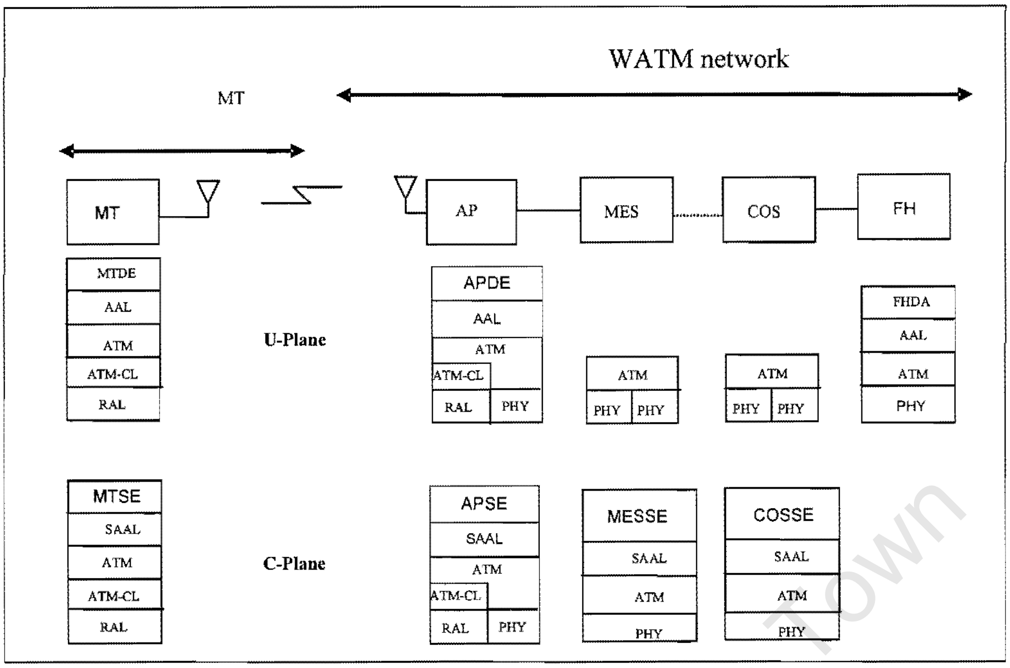


Fig. 4.2 Control and User Plane Protocol Layering

4.4 Handover Signaling Entity Design considerations

The Handover Signaling Entities can be developed to have certain general features that are expected of any good and efficient design [42]. These features are as discussed below:

- **Component-based development:** The signaling entity should be developed using a component-based approach as against a pure procedure-based approach. The component-based approach is attractive in that it is easier to develop and that it facilitates reuse. Also, faults and bugs are easier to locate in case of a component based approach. The various components of the handover signaling entities are discussed in detail in the next section.
- **Generic Routines:** The signaling entity can be written so as to have separate generic routines to implement data structures like trees, link-lists, etc. These routines are generic in that they can be reused with very little or no modification. Hence, reusability of components across different entity implementations is possible, if kept in mind initially.
- **Ease of Porting:** The signaling entity should be implemented so as to be processor as well as Operating System independent. If coding in "C" ("C" is the most commonly used programming language for signaling implementations), ANSI C can be used to ensure

ease of portability. Operating System related calls should always be made through macros, which can all be clubbed in one header file. This file can be named appropriately to indicate that this (and only this one) is to be touched for porting.

4.5 Handover Signaling Entity Architecture Design

A high level architecture of the handover signaling entities with various components is discussed in this section. The methodology of the architecture is structured around a component based design. This ensures that any future changes can be confined to small sections of the overall design.

In order to process a handoff request, the various control entities needs to support a set of functions that will facilitate the handoff request messages. These functions are divided into several modules in order to implement the desired functionality, as specified by the handover signaling protocol. This section presents the high level architecture design of the signaling entities at the MT, AP, MES and COS as described in section 4.2. For a better understanding of the flow of information within these entities, the reader is advised to consult Appendix D.

4.5.1 MTSE Architecture Design

Fig. 4.3 is an image showing the software modules in the MTSE software design and illustrating the interaction between the various modules. Within this figure, thick hollow arrows represent user data paths, thin arrows represent control paths for inter-module communication, and the thick solid arrows represent handover signaling paths.

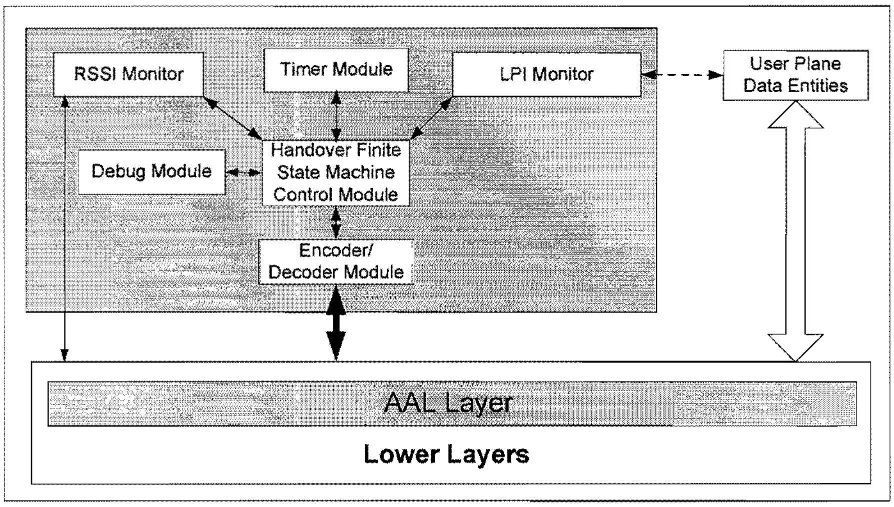


Fig. 4.3 High level software architecture of the MTSE

- **RSSI Module**: The MTSE should be able to monitor the radio signal strength in order to decide when a handover should be initiated. The MTSE should therefore incorporate a Radio Signal Strength Indicator (RSSI). The RSSI Module should have access to the radio physical layer in order to monitor its quality. If the radio signal quality drops below a set threshold value, the RSSI module should generate an interrupt to initiate a handover operation. The interrupt signal should be directed to the Encoder/Decoder Module for processing. This module is called the Radio Link Monitor.
- **Encoder/Decoder Module**: The handover-signaling entities exchange messages that have well defined formats. These messages can be formed at the time of transmission by an Encoder. The Encoder takes appropriate parameters as input and forms these well-defined signaling messages. During reception, a Decoder (also called Message parser) is used. The Decoder does the validation and extraction of all the messages received from the peer entity. If the validation is successful, appropriate action is taken, else error handler is called.
- **Timer Module**: High-resolution timers should be included in the implementation. These timers should be used to provide time stamping of handover signaling messages in order to allow performance monitoring of the handover signaling protocols. A time stamping module should also be implemented in the user data entities in order to monitor the flow of user data during a handover.

- **LPI Monitor**: This module performs management plane functionality in that it encompasses relevant functions to interact and coordinate between user and control plane activities. The LPI Monitor should have direct access to the user data paths. This is necessary because signaling information is inserted in-band into the data paths when a LPI message is required. This module is also responsible for generating the LPI cell and monitoring the user data paths for signaling messages (LPI packets). Furthermore, the LPI Monitor instructs the MTDE to buffer the uplink and downlink data as required by the handover operation.
- **Handover State Machine Control Module:** This is the most important part of the entity and is responsible for maintaining the state of all the protocol managed by SVC entity. The FSM module can be organized as a collection of state-event functions. Based upon the type of event and the current state of SVC entity, a function can be invoked using function tables. This module also ensures that all the different modules can operate independently of each other by providing a suitable common interface between these different modules. The Handover State Machine determines the type of handover protocol implemented, i.e. PE or PR Scheme.
- **Debug:** The Debug module (also called TRACE module) is responsible for printing of debugging traces (i.e., printf statements which track the flow of execution). Flexibility should be provided so that all of these can either be entirely compiled out for a faster and more-efficient code, or the level of details provided by these statements can be changed on a per-module basis. Both preprocessor and runtime options can be provided.

4.5.2 APSE Architecture Design

The APSE consists of the modules as illustrated in Fig. 4.4.

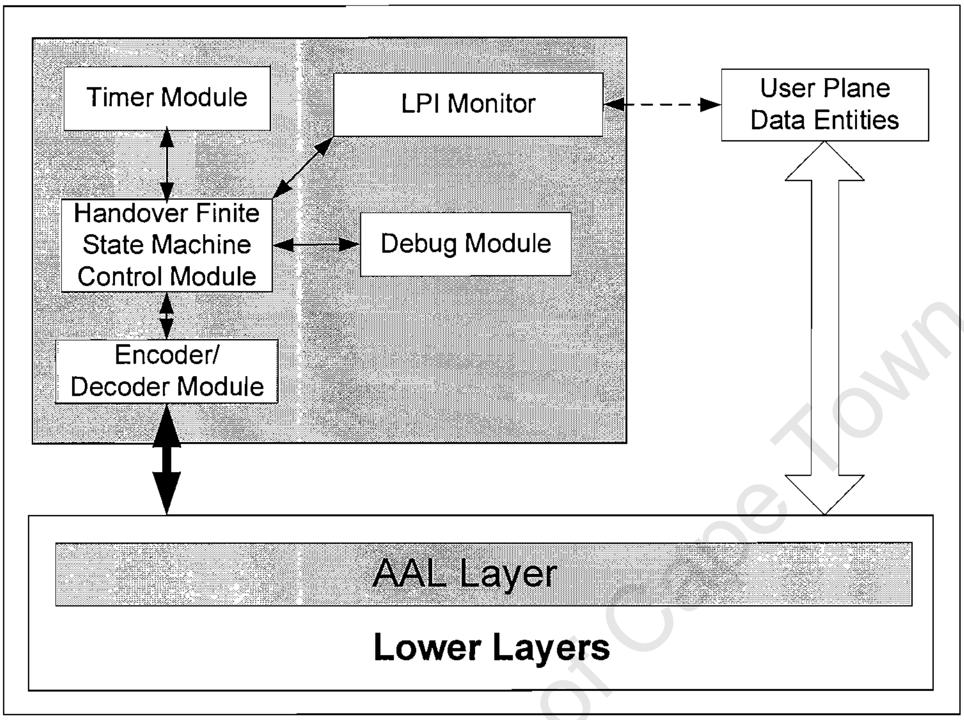


Fig. 4.4 High level software architecture of the APSE

The functions of these modules are the same as for the MTSE. The Finite State machine module for the AP determines whether the AP functions as the old or the new AP, e.g. when the AP functions as the old AP, the LPI monitor would not be incorporated in the handover operation.

4.5.3 MESSE Architecture Design

Fig. 4.5 is an image showing the software modules in the MESSE software design and illustrating the interaction between the various modules.

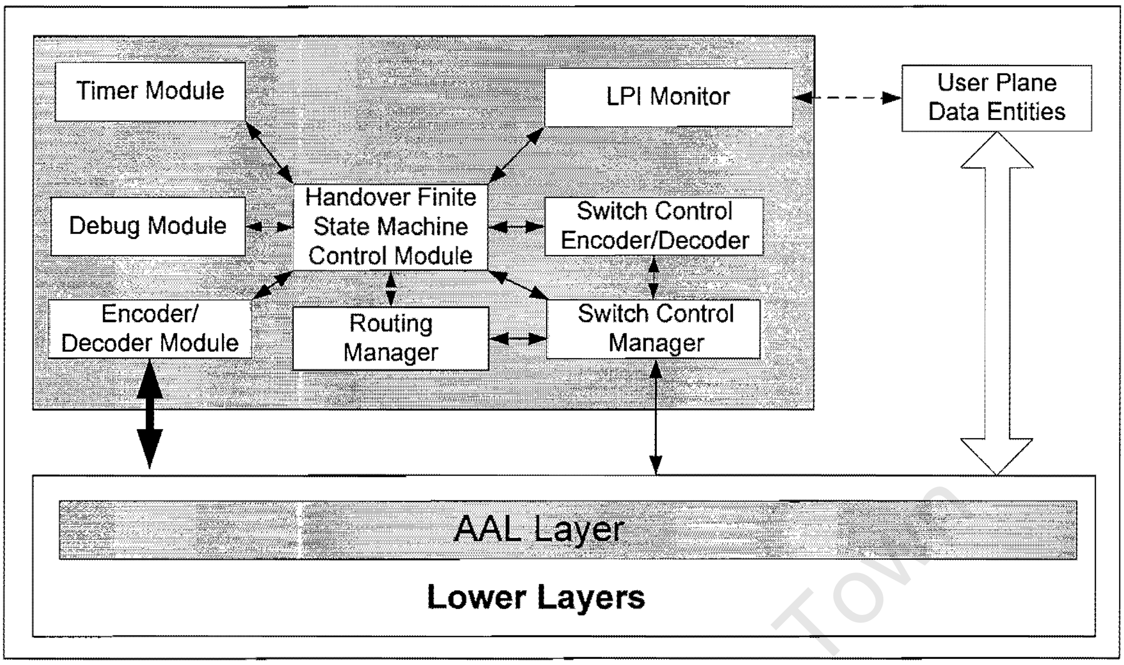


Fig. 4.5 High level software architecture of the MESSE

Again, the functions of the modules that are common to the MTSE and MESSE are identical. However, the following modules are also needed in the MESSE that provide access to the ATM switch hardware to facilitate handover of the users VCs when requested during handover. The extra modules should not be embedded in the switch hardware. The extra components include:

- **Switch Control Encoder/Decoder**: This module is similar to the handover signal encoder/decoder module, however, it is only invoked when switching of the downlink or uplink data paths are requested. This module encodes/decodes messages into formats that the Switch Control Manager can recognize.
- **Switch Control Manager**: The ATM switch is controlled by a Switch Control Manager (SCM). The SCM configures the switch hardware to route incoming cells to the appropriate outgoing links by modifying tables within the switch, thus establishing connections. The SCM in conjunction with the routing control entity within the switch, manages the switching resources, and is responsible for setting up or tearing down connections.

* **Routing Manager**: The Routing manager maintains routing information and handles all routing queries when a handover is requested. This is accomplished by locating switch resources that can satisfy the requirements of each handover connection request.

In order to realize the mobility functions required by a fixed ATM switch, the mobility extensions for handover should be implemented in an independent unit external to the switch. Therefore, the requirements for the MES Switch Control Manager are:

* Control Manager should be completely separate from the ATM switch
* Switch Control Manager interface to switch should be through ATM cells
* Low Level Script interface should exist to the Switch Control Software for programming handover control functions.

To a typical ATM switch architecture, the following additions are added externally to the switch to implement the MES. This is illustrated on Fig. 4.6.

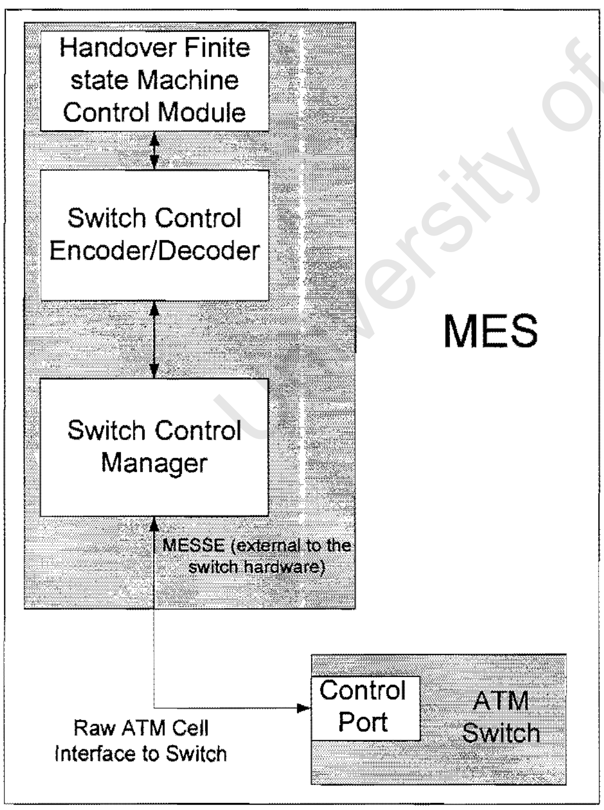


Fig. 4.6 MES Software integration

The switch control manager should have direct access to the switch routing tables via a pre-established VPI/VCI to a control port on the switch. The Handover Finite State Machine determines whether the MESSE functions as the old or the new MES.

4.5.4 COSSE Architecture Design

The COS Signaling Entity software design is identical to the MESSE. The only distinction between the two is contained in the Handover Finite State Machine. Thus, depending on the defined states in the entity, it will either operate as a MES or a COS.

4.6 Handover Signaling and Data flow between Entities

Having introduced the various entities involved in the handover process, we now take a closer look at the signaling and data flow between them. The interaction of the various entities involved in the handover process is illustrated in Fig. 4.7. The diagram shows the software entities and layout in the software design system. The diagram also shows which entities will be implemented in the MES, the APs and the MT. In the diagram, thin dashed lines represent signaling data paths while the thick solid lines represent user data paths. The Switch Control Encoder/Decoder and the Switch Control Manager is part of the MESSE, however, it is shown outside the MESSE in Fig. 4.7, in order to clarify its operations and interface with the switch hardware.

For a handover operation to commence, a signaling VC (SVC) needs to exist between the following modules:

- The old APSE and the old MESSE
- The new APSE and the new MESSE
- The old MESSE and the new MESSE, for the PE Scheme
- The old MESSE and the COSSE, for the PR Scheme
- The new MESSE and the COSSE, for the PR Scheme
- The MTSE and the old APSE
- The MTSE and the new APSE

To facilitate communication between the Switch Control Manager and the internal ATM switch routing tables, a control VC also needs to exist between them.

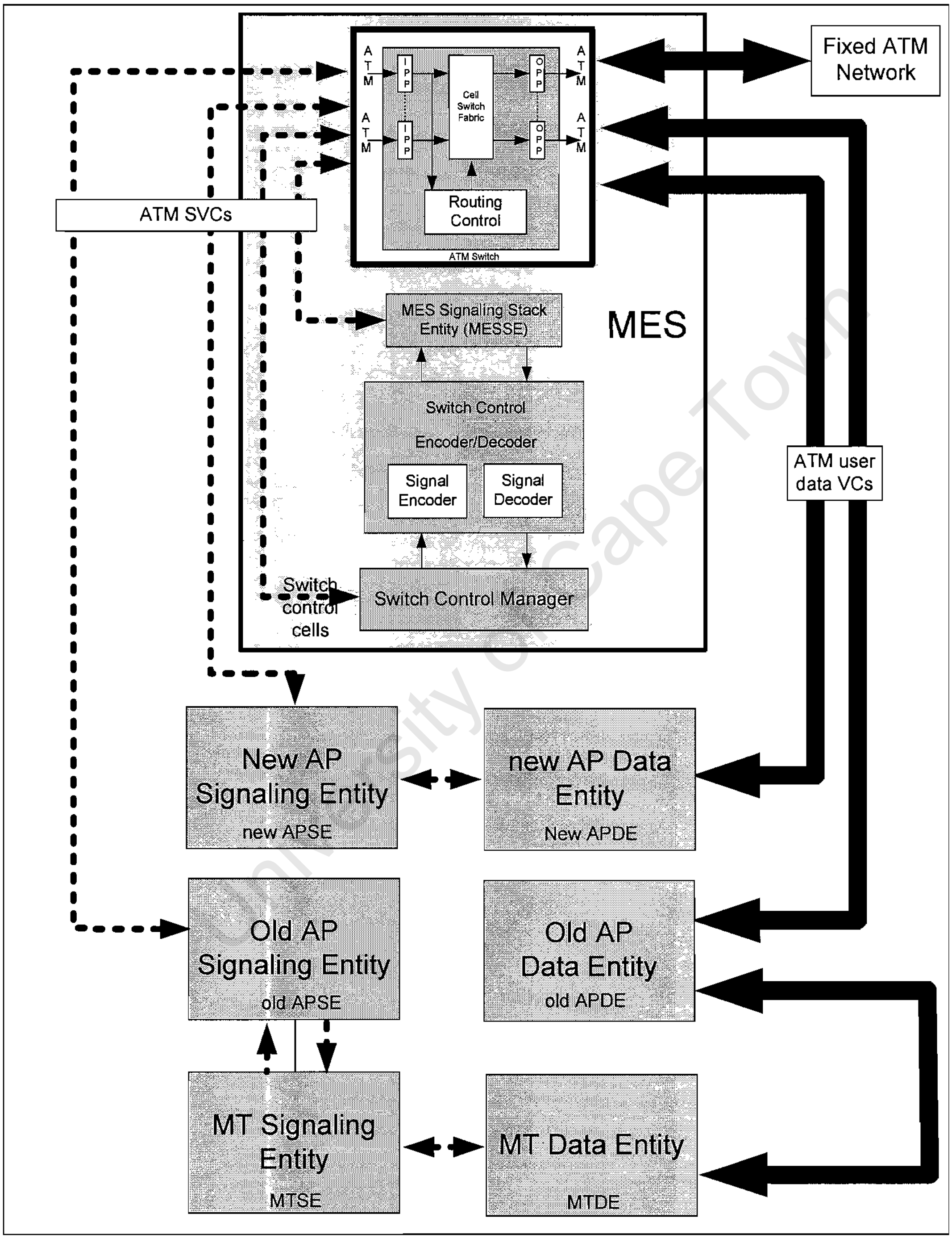


Fig. 4.7 Signaling and Data Flow Diagram

The user and signaling paths share the physical, virtual circuit (VC) and virtual path (VP) transmission facilities from the ATM network. The handover signaling services are supported over logical signaling links established by pre-assigned permanent VCs within the ATM network. Hence, ATM cells comprising the signaling messages are sent on separate VC's that are provided by the network for signaling purposes.

The signaling VC's should have first priority for transmission at intermediate network nodes.

The design assumes that a data VC exits end-to-end between the MT and its FH. The existing data path will be rerouted in the ATM network according to the applicable handover scheme utilized.

In order to simulate more than one hop between the AP and the COS, path looping within the same switch can be easily accomplished with this architecture. The signaling path is rerouted via the same switch twice to emulate two hops, three times to emulate three hops, etc.

Chapter 5

WATM and signaling framework

5.1 Introduction

The previous two chapters described the handover signaling protocols and the design of the handover signaling entities. This chapter focuses on the architecture of the WATM network and signaling entities required to process the handover protocols. This architecture provides a platform on which to test and examine the analysis performed in chapter 3.

5.2 WATM Evaluation Platform

The test bed consists of hardware and software components. The architecture used in the evaluation platform is illustrated in Fig. 5.1. It consists of five workstations and a Washington University Gigabit ATM Switch (WUGS, [35, 36]).

Four workstations serve as a FH, MT and APs. The WUGS together with one workstation running the MESSE software forms the MES. The FH, MT and MESSE workstation runs the Linux 2.4.18 Operating System, while the two APs run the NetBSD 1.4.1 Operating System.

The wireless link between the MT and APs are emulated via an Ethernet network. There are no wireless ATM cards available for experimentation. However, this is not an obstacle, because the transmission characteristics of a wireless link can be compared to the transmission characteristics of a wired (Ethernet) link in the following manner:

When a handover signaling message is transmitted in the wireless medium, the only important consideration for handover evaluation purposes is the propagation time of the signaling message in the wireless link. The fact there exist no physical connection between the MT and AP is of no consequence for handover evaluation purposes. The same argument is valid for a wired link.

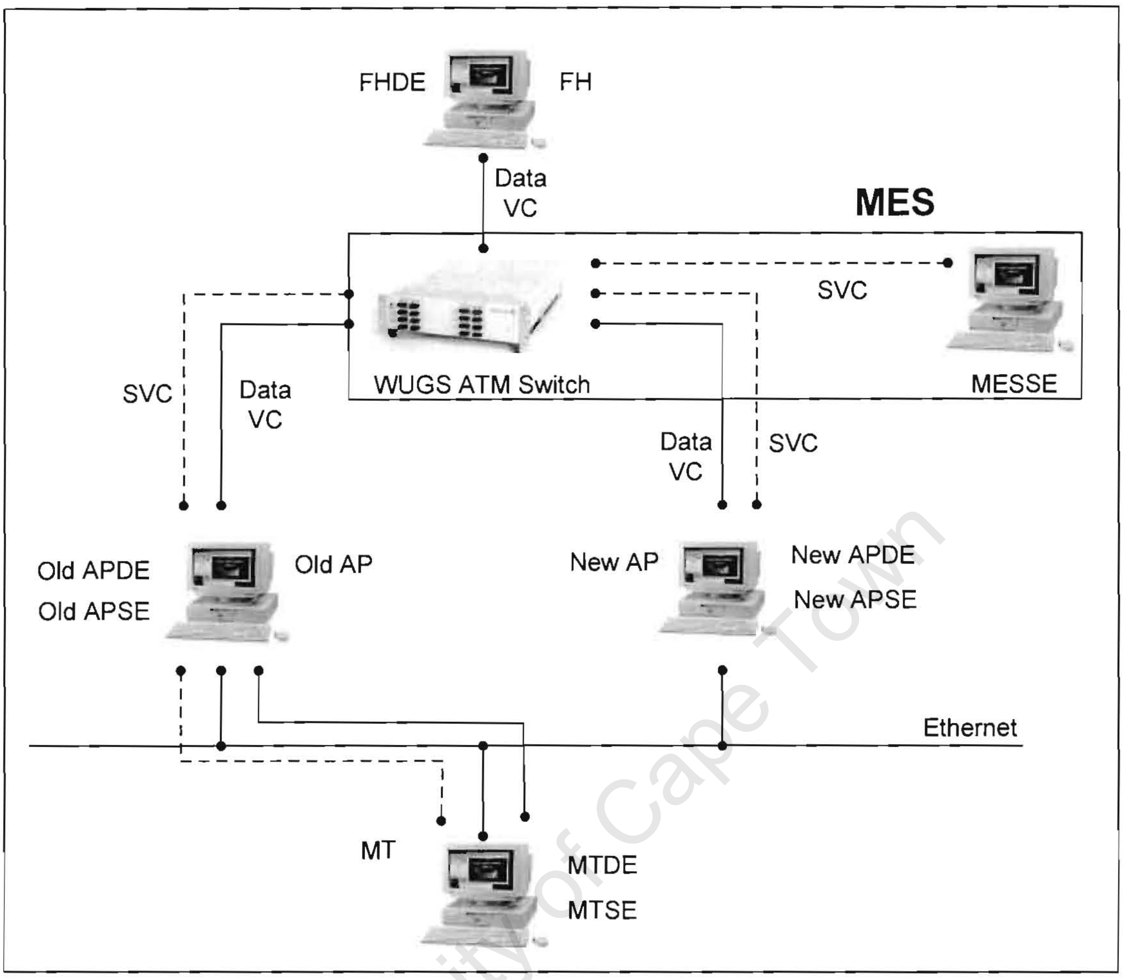


Fig. 5.1 Architecture for the WATM network

Past experiments were performed to evaluate the performance of handoff schemes using both Wavelan wireless Ethernet cards and conventional wired Ethernet [7]. It was found that the results of the performance evaluation are similar using both configurations. However, packet losses using the wireless configuration are slightly higher. On further investigation, it was found that the reason for the higher packet losses in the wireless link was due to occasional burst losses suffered in the wireless medium. It is for this reason that the choice of using Ethernet is revalidated to eliminate the wireless specific effects and concentrate on the losses due to handover and connection rerouting in isolation.

The protocol stack for the architecture of Fig. 5.1 is illustrated in Fig. 5.2.

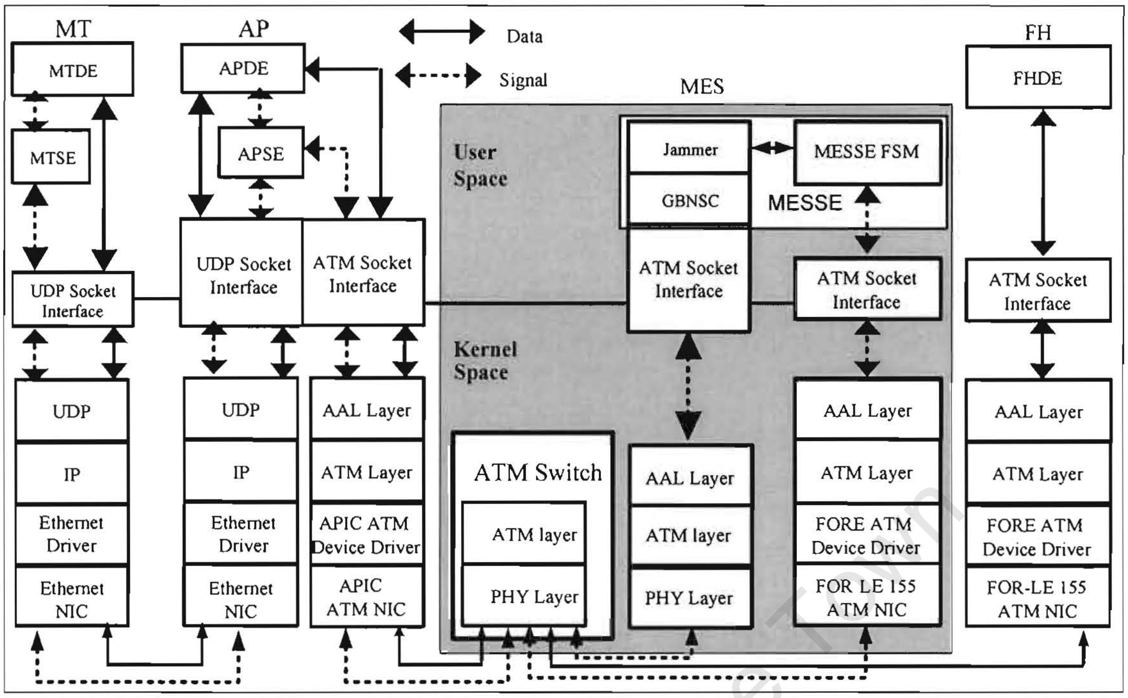


Fig. 5.2 Protocol Stack for WATM network architecture

The network protocol stack at the FH and the MT are different. The FH uses the ATM Adaptation Layer (AAL5) protocol over ATM whereas at the MT, the UDP/IP over Ethernet is used. The APs performs the bridging function between the different protocol stacks.

The performance evaluation of the handover schemes with the WATM architecture illustrated in Fig. 5.1 will apply to an end-to-end ATM system with wireless endpoints as well. The difference with the evaluation platform architecture and an end-to-end wireless ATM system is that, in the evaluation platform, the APs are responsible for reassembling ATM cells into UDP packets. In an end-to-end ATM system, the MT would reassemble the ATM cells. However, cell loss during handover results in the loss of one or more application packets in both systems, regardless of where reassembly is performed.

For a step by step explanation of the construction of the WATM test bed with its protocol stack, the reader is advised to consult Appendix H.

The two APs are connected to the WUGS via multimode fiber. Both APs are equipped with an ATM Port Interconnect Chip (APIC – ATM NIC developed at Washington University) NIC. The FH and MESSE workstation are also connected to the WUGS via multimode fiber. However, theses two PCs are equipped with ATM NIC Efficient Networks Fore 155 LE that is connected with dual OC-3 line cards at the switch ports. All devices also contain an Ethernet NIC.

The only reason the APIC NIC was used in the APs is due to a limited no. of Efficient Networks Fore 155 LE NICs being available for experimentation. The author would have preferred using the Fore NICs due to the extensive support being available for ATM-on-Linux. The use of the APIC NICs is also the reason why NetBSD 1.4.1 was used in the APs, as opposed to the Linux kernel. The ATM device drivers for the APIC NIC were specifically written for the NetBSD kernel, and no support is available for Linux. However, the choice of ATM NIC used only facilitates the ease of implementation and has no effect on the performance of the handover schemes.

The Switch Control Manager uses the Gigabit Network Switch Controller (GBNSC) interface to the WUGS to modify the internal switch routing tables [37]. The Switch Control Encoder/Decoder uses the Jammer interface to encode/decode switch control message in the format that GBNSC understands [39]. The Jammer interface is also used to obtain switch routing information for handover purposes.

5.2.1 The Switch Control Manager - GBNSC

The purpose of this software is to control the WUGS and hide hardware details as much as possible. GBNSC monitors the state of the switch and provides access to all hardware details. The WUGS has no processing engine, thus, a workstation running GBNSC controls the switch. Access to the switch is through ATM cells transmitted by the controller. These special cells, called control cells, have special formats defined and are sent on VPI 0 VCI 32. The switch's internal routing tables and maintenance registers are modified and monitored by the control cells.

5.2.2 The Switch Control Encoder/Decoder - Jammer

Jammer is a script-based client utility used to access all the bits in the WUGS tables and registers. It connects to GBNSC through a TCP/IP socket (not shown in Fig. 5.2) and issues pre-defined commands to ping the switch, read or write routing tables, read maintenance registers, or reset/clear the switch. Users can create Jammer scripts to automate routing table programming. Appendix B presents a script used with implementation.

5.3 Capturing Time and Sequence numbers

Every data and signaling packet transmitted in the experimental WATM test bed included a sequence number and the time at which the packet was sent. The time and sequence of the data and signaling packets were logged at the MT, the APs and the MES. This was important in order to track the sequence of the packets during the handoff and to be able to quantify delays as the packets traversed the network. The Network Timing Protocol was used to provide synchronization between the various network elements [54]. Network Time Protocol (NTP) is a distributed computer clock synchronization protocol. NTP can be used in various modes. NTP is widely used in the classic client-server mode with a hierarchy built in to reduce network traffic and latency. NTP can also be used in symmetric mode by isolated networks, such as a peer to peer network. Finally, NTP can operate in a broadcast mode if there are a large number of clients involved.

The standard time used by most nations of the world is Universal Coordinated Time (UTC), formerly known as Greenwich Mean Time (GMT). NTP uses UTC to synchronize "primary" servers via radio, satellite receiver or modem. These primary servers then adjust the clocks of secondary servers/clients. In order to correctly adjust clocks of secondary servers over a LAN or WAN, a time offset of the server clock relative to the client clock is computed by the client running NTP. In existence today, there are 79 public primary servers synchronized directly to UTC. There are over 100 public secondary servers synchronized to the primary servers and providing synchronization to more than 100,000 clients and servers in the Internet. Additionally, there are an unknown number of private servers utilizing NTP. The general model for discovering the clock offset starts with a server sending a message that includes its current clock value to the client, which could be another server or workstation. The client records its own current clock value upon arrival of the message. For accuracy, the client has to measure the server-client propagation delay. NTP measures the total roundtrip delay and assumes the propagation times are statistically equal in each direction.

Clock errors are due to variation in network delay and latencies in computer hardware and software (jitter), as well as clock oscillator instability (wander). According to NTP documentation, NTP in the majority of cases can keep clock synchronization within a few milliseconds on LANs [54]. This performance is acceptable for the WATM test bed.

5.4 User Plane Data Entity Implementations

Fig. 5.3 presents a graphical illustration of the FH, AP and MT Data Entities.

5.4.1 MT Data Entity (MTDE) Implementation

The MTDE acts as a UDP/IP receiver and transmitter. It receives the data from the UDP socket and puts it in a data storage buffer. The buffered data will then be written to a file. The MT is also capable of transmitting data in the same way as the Fixed Host Data Entity (FHDE).

5.4.2 Fixed Host Data Entity (FHDE) Implementation

The FHDE acts as an AAL5 transmitter and receiver. In the transmit mode it reads data from a file and into a buffer. The buffered data is then sent to a socket for transmission. The FHDE is capable of reading and writing data. A timer interrupt function controls the rate at which data is sent.

5.4.3 Access Point Data Entity (APDE) Implementation

The function of the APDE is to act as a gateway to interconnect the ATM network and the Ethernet Network.

The method used in the Access Point is to extract the information from the ATM socket and put the data into the Ethernet socket or vice verse if the data is sending from the Ethernet side. The Access Point is designed to keep listening to the source (ATM and Ethernet).

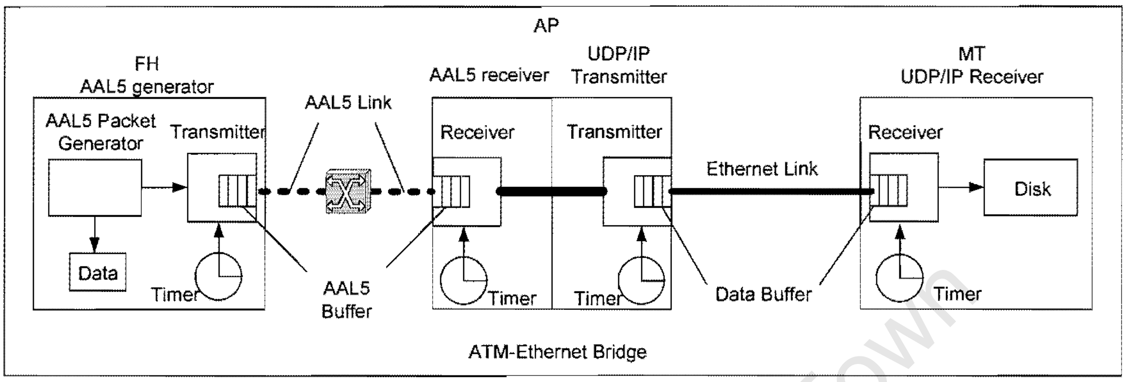


Fig. 5.3 Test bed data entity software components

5.5 Experimental Configurations

The WATM architecture as illustrated in Fig. 5.1 was used to evaluate the performance of both the PE and the PR Schemes.

In the first configuration, the test bed was set up to study the performance of the PR scheme. In order to emulate more than one hop between the APs and the COS, the configuration shown in Fig. 5.4 was set up.

The test bed makes use of path looping within the ATM switch. Logically, Fig. 5.4 can be illustrated with Fig. 5.5.

Using this manner, the signaling messages are not aware that only one switch is present. As far as it is concerned, it had to travel from the old AP to the COS via two ATM Links. Handover signaling messages occur sequentially, therefore the multiple processes (old MESSE, new MESSE and COSSE) running on the same workstation do not compete for system resources.

The User data paths can be setup in the same manner. The complete listing of handover signaling VC's used in the test bed is given in Appendix E.

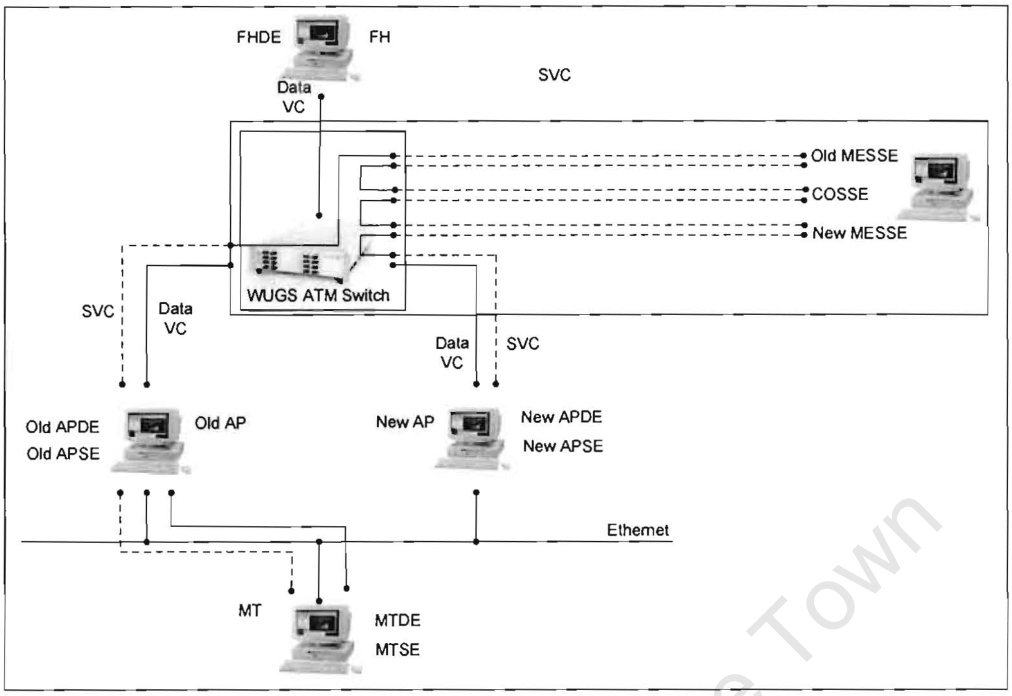


Fig 5.4 Emulating multiple hops with the WATM architecture

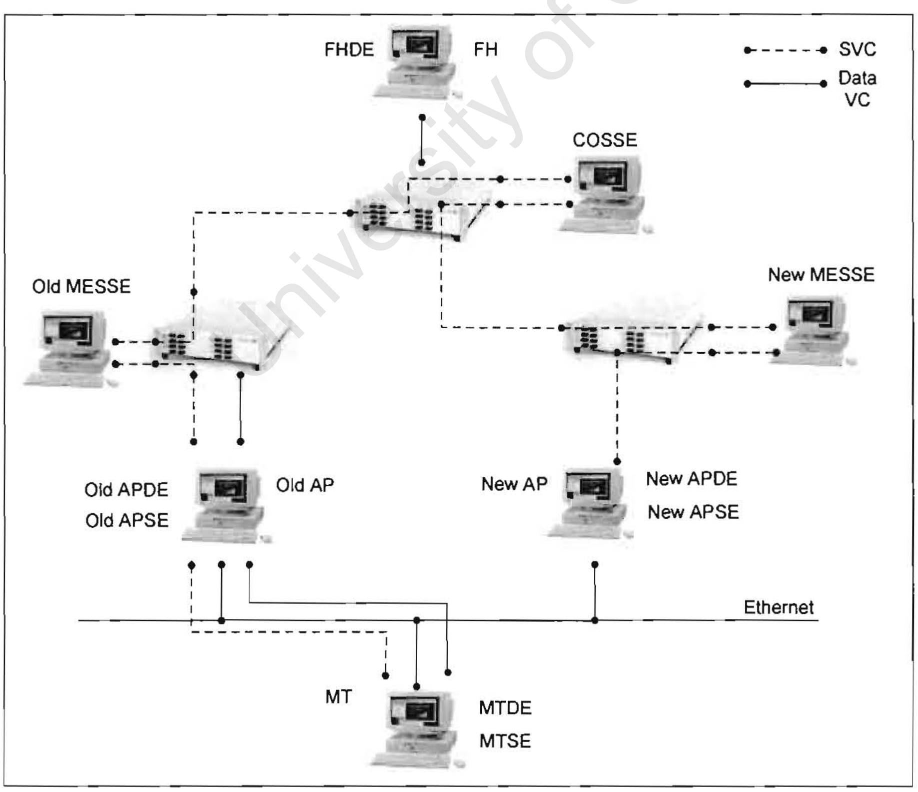


Fig. 5.5 Logical Diagram of WATM architecture for PR scheme, with multiple hops

In order to emulate intermediate switches between the MES and the COS, an Intermediate Signaling Entity (ISE) were running on the same workstation as the old MESSE, new MESSE, and the COSSE. The design of the ISE is not a new one, it is the same as a COSSE, however, the ISE simply processes the handover signaling messages and passes it to the next switch in the signaling path without requesting the switching of the user data VCs from the switch.

The second test bed configuration was setup to study the performance of the PE scheme. This test bed configuration is similar to the configuration shown in Fig. 5.4. However, there is no COSSE and a direct link (PVC) exists between the old MES and the new MES. Therefore, the logical test bed configuration that exists is illustrated in Fig. 5.6.

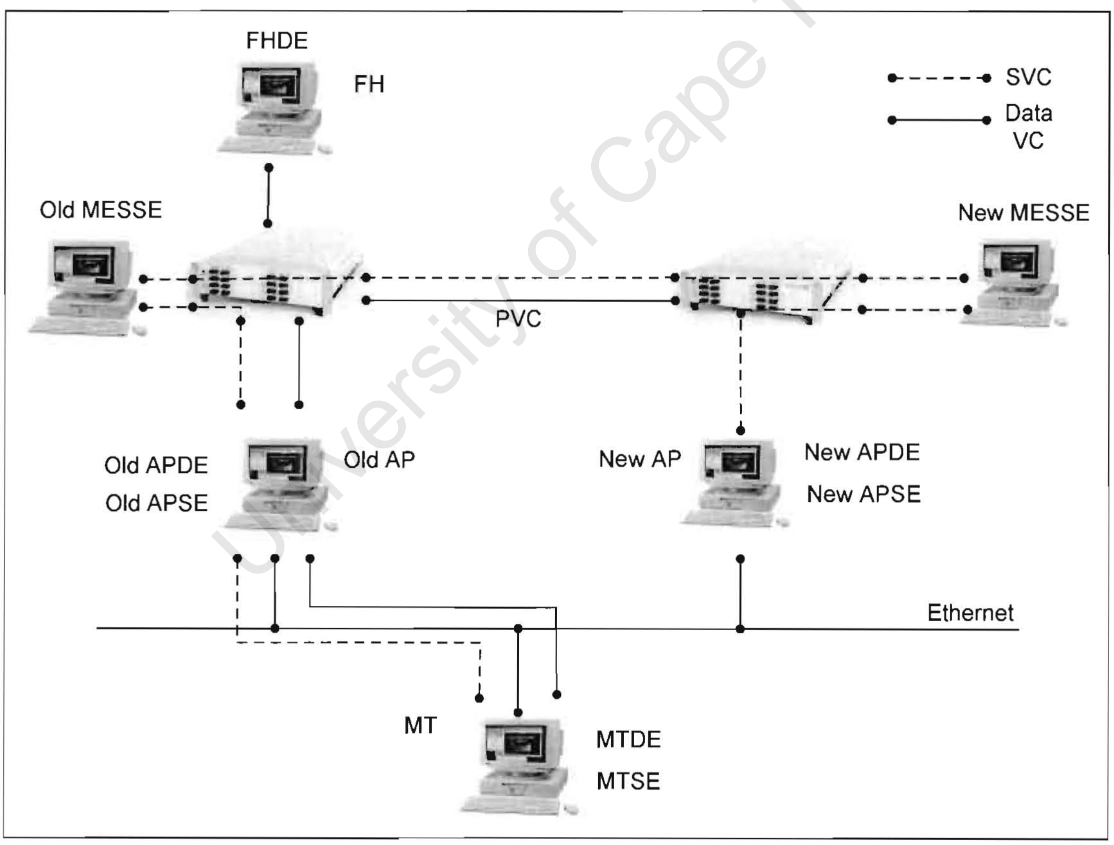


Fig. 5.6 Logical Diagram of WATM architecture for PE scheme

5.6 User data traffic sources

This section describes the characteristics of the traffic sources used to evaluate the performance of the PR and PE scheme. Traffic sources used to evaluate the performance characteristics of the two schemes were delay sensitive Constant Bit Rate (CBR) traffic.

Audio traffic can be characterized by its timing component. Every T milliseconds, a 160-byte packet will be transmitted by the source. For T = 20, this corresponds to telephone quality audio of 64 kb/s. In a typical audio conversation, there will be silence periods during which no packets will be transmitted. If this occurs during a handoff, there would be no disruption of traffic. In the experimental setup, the traffic generator transmits continuously in order to detect all potential disruption due to handover.

The Video traffic can be characterized by a Constant Bit Rate source sending packets every 30ms. The size of the packet and hence the bandwidth required depends on the particular video format used. Table 5.1 lists common live-capture video formats, their approximate bandwidth requirements and the typical application usage.

Video format	Bandwidth	Usage
MPEG4	28 – 768 kbps	Good image quality and optimized for low data rates
H.261	128 – 1000 kbps	Low motion applications such as video conferencing
MPEG1	500 – 1500 kbps	Motion picture or comparable VHS-quality
Indeo 4.1	1000 – 1500 kbps	Picture quality video applications
MPEG2 Half D1	2 – 3.5 Mbps	Half D1 offers better quality than MPEG2 Full D1
MPEG2 Full D1	3 – 6 Mbps	DVD-quality video and optimized for higher data rates

Table 5.1 Video Formats

By varying the packet size, any of the above video formats can be emulated. However, instead of transmitting real audio and video, experiments were performed by making use of a continuous traffic (Constant Bit Rate – CBR) source at the MT and FH, so that all potential disruptions during a handoff can be detected. The maximum throughput reached by the AAL5 traffic generator was 1.5 Mbytes/s. This is sufficient to support the MPEG1 video format.

5.7 Evaluation Platform Discussion

Handover requests were triggered periodically from the MT with an interrupt signal, with sufficient time between successive handoff requests to ensure that any connection reroute operation triggered by a previous handoff is complete before the next handoff request is issued.

The position of the COS is manually allocated and the intermediate switches does not perform a COS location algorithm. Further, no admission control tests were performed at intermediate switches. Therefore, it is expected that the latencies present in the handover operation would be significantly less than in a real ATM network.

All handover performance measurements were taken and averaged over 100 handoff attempts.

Chapter 6

Framework Evaluation and Results

6.1 Introduction

This chapter discusses the evaluation of the WATM architecture and handover signaling protocols. The first section discusses the performance evaluation of the MES, where particular focus is placed on the MES ability to switch packets correctly during a handover. The rest of this chapter is devoted to discussing the results of the tests performed to evaluate the operational aspects of the WATM architecture and handover signaling protocols.

The performance measurements to be considered with the experimental setup for the handover schemes are (i) Service disruption time, (ii) Handoff completion time, (iii) Buffer requirements at MT, and (iv) Buffer requirements at AP.

6.2 Performance evaluation of the MES

6.2.1 End-to-end Transmission Performance Evaluation (FH to MT)

To analyze the correct operation of the transmitter at the FH and receiver at the MT, the following requirements had to be met:

- Each packet received would have to match the transmitted packet exactly.
- Packets need to arrive in the correct order in which they were sent.
- When a handover occur, the packet should arrive in the correct order and without errors.

Packet 1, No. of bytes sent: 48
Contents 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7
Packet 2, No. of bytes sent: 96
Contents 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Packet 3, No. of bytes sent: 144
Contents 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
Packet 4, No. of bytes sent: 192
Contents 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9
Packet 5, No. of bytes sent: 240
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3
Packet 6, No. of bytes sent: 288
Contents 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7
Packet 7, No. of bytes sent: 336
Contents 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Packet 8, No. of bytes sent: 384
Contents 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
Packet 9, No. of bytes sent: 432
Contents 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9
Packet 10, No. of bytes sent: 480
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3
Packet 11, No. of bytes sent: 528
Contents 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7
Packet 12, No. of bytes sent: 576
Contents 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Packet 13, No. of bytes sent: 624
Contents 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5

Fig. 6.1 AAL5 Packet log file showing 13 Packets, each of length 48 bytes

The method used for comparing streams of packets to determine their accuracy, was to log all packets that were generated at the transmitter, as well as packets received at the receiver. A typical log file of packets is shown in Fig. 6.1.

Each transmitted packet also contained a header in which the packet number and the total number of bytes sent. This is used to compare the transmitted packets with the received packets. With the MES operating in a static mode (no handoff, normal ATM switch), the transmitted log file and received log file were identical. Hence, transmitted packets and the received packets always arrived in order and without errors at the receiver.

6.2.2 MES Performance Evaluation

To analyze the correct operation of the Mobility Enhanced ATM Switch during a handoff, the following tests were performed. With this test we are only concerned with testing the MES ability to correctly reroute a connection during a handover, and are not interested in the handover protocols yet. Thus, tests were only performed for data flow in one direction.

Packets were sent from the transmitter at the FH and also sent to a log file, Fig. 6.2.

```
Packet No. *1
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *1
Packet No. *2
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *2
Packet No. *3
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *3
Packet No. *4
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *4
Packet No. *5
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *5
Packet No. *6
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *6
Packet No. *7
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *7
Packet No. *8
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *8
Packet No. *9
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *9
Packet No. *10
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *10
Packet No. *11
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *11
Packet No. *12
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *12
Packet No. *13
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *13
```

Fig. 6.2 Packet log file at the transmitter (FH)

While packets are transmitted at the FH and received by the MT, a handover request is triggered at the MT. The received packets at the APs were also logged. At first data was sent via the old AP (indicated at the left of Fig. 6.3), and then a handover was invoked to the new AP (indicated to at the right of Fig. 6.3).

```
Base Station Aztec Now Active, packet no. *1
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *1
Base Station Aztec Now Active, packet no. *2
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *2
Base Station Aztec Now Active, packet no. *3
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *3
Base Station Aztec Now Active, packet no. *4
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *4
Base Station Aztec Now Active, packet no. *5
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *5
Base Station Aztec Now Active, packet no. *6
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *6
```

```
Access Point Inca Now Active, packet no. *7
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *7
Access Point Inca Now Active, packet no. *8
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *8
Access Point Inca Now Active, packet no. *9
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *9
Access Point Inca Now Active, packet no. *10
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *10
Access Point Inca Now Active, packet no. *11
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *11
Access Point Inca Now Active, packet no. *12
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *12
```

Fig. 6.3 Packet log file at both APs (during a handoff)

```
Packet No. *1
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *1
Packet No. *2
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *2
Packet No. *3
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *3
Packet No. *4
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *4
Packet No. *5
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *5
Packet No. *6
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *6
Packet No. *7
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *7
Packet No. *8
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *8
Packet No. *9
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *9
Packet No. *10
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *10
Packet No. *11
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *11
Packet No. *12
Contents 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 *12
Packet No. *13
Contents 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 *13
```


Fig. 6.4 Packet log file at the MT

Fig. 6.3 shows that just after the MES transmitted packet no. 6 to the old AP, the internal switch routing tables were modified to transmit packets to the new AP. As can be seen from Fig. 6.3., packet no. 7 arrived at the new AP after the handoff. No data is lost and no errors occurred in the packets. By considering Fig. 6.4, it can be seen that the packets also arrived at the MT receiver in order and with no errors.

When the bandwidth requirement of an application is specified, the value is normally given as the number of bits transmitted during a given period of time, e.g. bps (bits per second). However, this does not say much about the manner in which data is transmitted. Consider the following scenario:

A transmitter generates packets of length 1 Mb and transmits it every second. Therefore the transmit speed is 1 Mbps. A transmitter generating smaller packets will have to transmit it faster in order to retain the same throughput. This is a very important consideration in order to determine the maximum VC bandwidth that can be correctly supported by the MES during a handover.

Define T as the time interval between packet transmissions. Then the transmitter throughput is:

$$\frac{packet\_size}{T}$$

This packet size, as mentioned above could be larger than the ATM cell size of 53 bytes. It is the responsibility of the AAL layer to segment the larger packets into ATM cells. AAL layer should segment the larger packets into ATM payload at a rate of:

$$\frac{53bytes \times T}{packet\_size} = \text{time interval between ATM cell transmissions.}$$

Consider the basic operation of an ATM Switch: An ATM cell is received across a link on a known VCI or VPI value. The switch looks up the connection value in a local translation table to determine the outgoing port (or ports) of the connection and the new VPI/VCI value of the connection on that link. The switch then retransmits the ATM cell on that outgoing link with the appropriate connection identifiers.

In order to ensure no loss of data in the MES during a handover, the time it takes to change the local translation tables should be less than the time interval between ATM cell transmissions.

The average time it takes the WUGS to perform an update of the local translation tables upon reception of a handover request by the Switch Control Manager is 200ns [36]. Since ATM cells are always 53bytes, the WUGS can correctly perform a handover operation for a maximum VC bandwidth of:

$$BW_{\max} = \frac{53bytes}{200ns} = 265000000bytes/s = 265Mbps$$

This speed was, however, not attainable with the AAL5 generator at the FH. The maximum throughput reached by the AAL5 generator was 1.5 Mbps.

6.3 Performance Evaluation of the handover schemes on the Wireless ATM Test Bed

The service disruption time, time taken to complete a handoff and extra buffering needed to avoid data loss due to rerouting are important performance issues to be considered as they affect the quality of service of the connections.

The two rerouting schemes discussed in chapter 3 have been implemented on the WATM test bed. The performance measurements considered are (i) Service disruption time, (ii) Handoff completion time, (iii) Buffer requirements at MT, and Buffer requirements at AP.

6.3.1 Performance Evaluation of the Partial Re-Establishment Scheme

6.3.1.1 Service Disruption Time

The service disruption time for the Partial Re-establishment scheme is shown in Fig. 6.5.

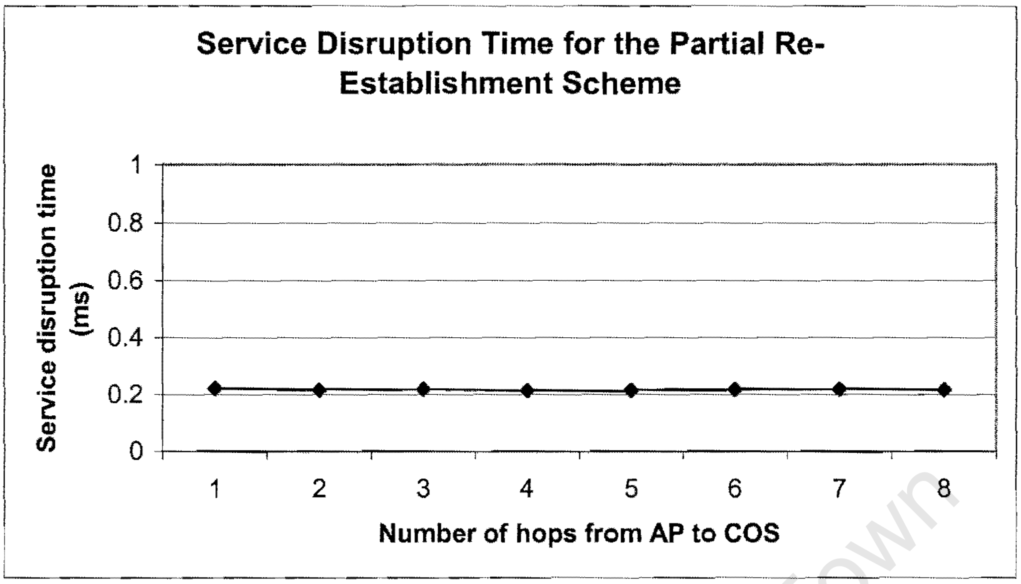


Fig. 6.5 Service Disruption Time for the Partial Re-establishment Scheme

One hundred handoffs were performed for each number of hops and the results were averaged to obtain the graph shown. Fig. 6.5 shows that the service disruption time for the Partial Re-establishment Scheme is constant with the increase in the number of hops between the COS and the Access Point. The reason for this is that the service disruption time will not be affected by multiple ATM switch handoff, since the disruption is determined only in the last hop between the AP and the MT. Recall from chapter 3 that the service disruption time, $T_{disrupt}$, 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 signaling messages at the switches and the transmission time of the first data packet. From the signaling sequence of the Partial Re-establishment Scheme it can be seen that the disruption is only dependent on messages 12 and 13 (Fig. 3.3 and Equation 3.32). Hence, for any number of hops the service disruption time is constant for the Partial Re-establishment Scheme. This experimental result confirms the analytical analysis as illustrated in Fig. 3.5.

6.3.1.2 Handoff Completion Time

The handoff completion time for the Partial Re-establishment scheme is shown in Fig. 6.6.

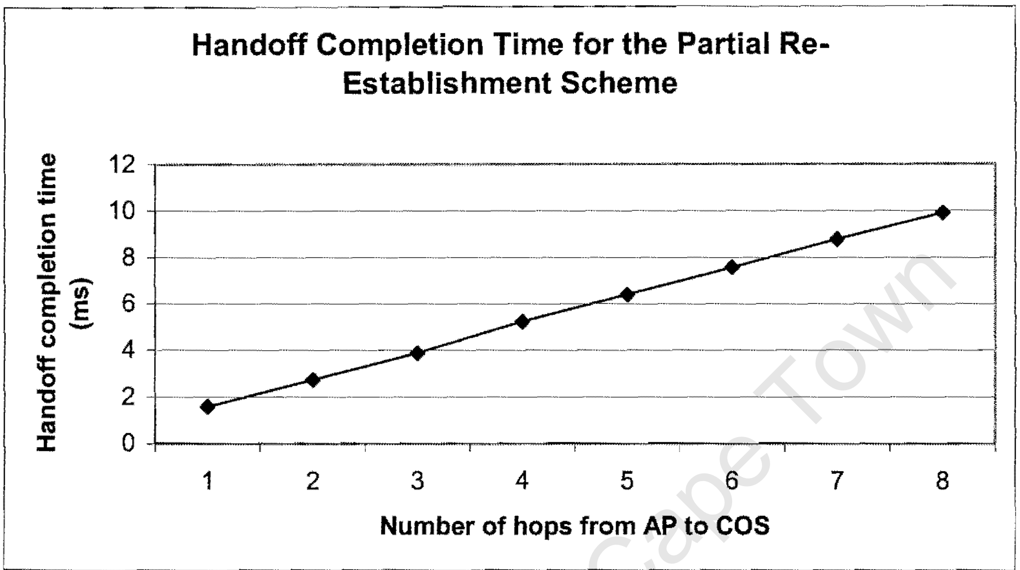


Fig. 6.6 Handoff completion time for the Partial Re-establishment scheme

The same procedures were performed as in the service disruption time experiment (see above) where one hundred handoffs were performed and the results averaged. Fig. 6.6 shows that the handoff completion time increases linearly with the increase in the number of hops between the Access Point and the COS. The reason for this is that a completely new path is setup from the COS to the MT, hence the connection setup time is directly proportional to the number of hops between the MT and the COS.

The handoff completion time, $T_{complete}$, 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 previous 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.

6.3.1.3 Buffer Requirements at the AP

The buffer requirements at the Access Point for down link data was found to be always zero.

Recall from chapter 3 that after the new AP sends out the Radio Resources Confirm message (RR_CONFIRM, message 5), it implements a buffer that waits for data on the new path. It then buffers all data until it transmits the Radio Channel Confirm message (RC_CONFIRM, message 13) to the MT.

After the COS receives the VC_CONFIRM message (9) from the new MES, it switches the newly arriving downlink data to the new AP.

The amount of buffering required at the new AP, is equal to the amount of data that is transmitted in the new path before the MT is connected to the new path. To understand why the buffering at the new AP is zero, one needs to look at how data is transmitted at the FH. This is a very important aspect that the analytical analysis of chapter 3 does not discuss or consider. This consideration was already introduced in section 6.2.2.

Recall from chapter 5 that the FH acts as an AAL5 transmitter, it reads data from a file and into a buffer. The buffered data is then sent to a socket for transmission. The FH is capable of reading and writing data. The timer function controls the rate at which data is sent.

In the AAL5 generator, the timer function control is implemented with an interrupt that will trigger the transmitter whenever a packet is ready. It was found that with the Linux and NetBSD operating systems, an interrupt could only be generated with a minimum of 10ms. Therefore, the minimum time between sending packets is 10ms.

Consider Fig. 6.5 and 6.6, it can be seen that the service disruption time is a constant 0.3ms and the largest handover completion time (corresponding to 8 hops) is about 10ms. Hence, by the time the AAL5 transmitter is ready to send the next packet, the service disruption time and the handover completion time is complete. Therefore, no data is yet transmitted in the new path before the handover or service disruption time is complete.

The analytical analysis assumes that data is transmitted in a continuous manner. It does not consider the fact that the data must first be read into memory (at the transmitter), then packed into data packets and sent out in a discreet manner. There is a limit to the time taken to perform these functions (packetization delay). The time between interrupts is a limit imposed by the operating system.

This theory was tested by adding a delay in the signaling and to check what effect it had on the buffer at the new AP. If the theory is correct, then by adding a 1s delay should result in a buffer size of:

$$BufferSize = \left(\frac{Time\_delay}{10ms}\right) * 48$$

Hence, for a 1s delay, the buffer requirements at the AP should be 4800 bytes. In the experimental setup, a delay of 1s was added to the handover signaling, and it was found that the buffer required at the AP is indeed 4800 bytes.

In order to solve the time limit problem imposed by the operating system on the interrupt, the handover signaling is slowed down in order for the operating system to keep up. In order to do this, the smallest time period at which the operating system interrupt can operate (i.e. 10ms), is declared to be 1 unit. 1 unit will correspond to 10 μs. Therefore; a real time interval of 1 ms will be represented by 100 units, or 1s. Hence, a thorough performance investigation can be conducted as the results can simply be scaled back to real time units if required.

In order to determine the buffer requirements at the AP, a delay was added to the handover signaling according to the method described above. The handover completion time of Fig. 6.6 was scaled according to the time unit. For the 1 hop example, the handover completion time was 1.56 ms. This corresponds to a time unit of:

$$1.56ms = \left(\frac{1.56^{-3}}{10^{-6}}\right) Units = 156 units$$

Hence, 1.56ms is scaled to 1.56 s. The same scaling procedure was done for the other handover completion times. This is shown in Table 6.1.

The buffer requirements at the AP with the scaled signaling delay is shown in Figure 6.7. The buffer requirement at the AP for downlink data is directly proportional to the handoff completion time. As the number of hops between the AP and COS increase, so does the buffer requirements at the AP in a linear manner. The reason for this is that as the hops increase, the time taken to reach the MT to confirm the handoff at the COS also increase. Hence, more data is transmitted in the new path before the MT is connected to the new AP.

Table 6.1 Handover Time normalization

Number of Hops	Handover completion Real-Time	Handover completion Time Units	Handover completion Scaled Time
1	1.56 ms	156 units	1.56 s
2	2.71 ms	271 units	2.71 s
3	3.86 ms	386 units	3.86 s
4	5.21 ms	521 units	5.21 s
5	6.38 ms	638 units	6.38 s
6	7.55 ms	755 units	7.55 s
7	8.76 ms	876 units	8.76 s
8	9.89 ms	989 units	9.89 s

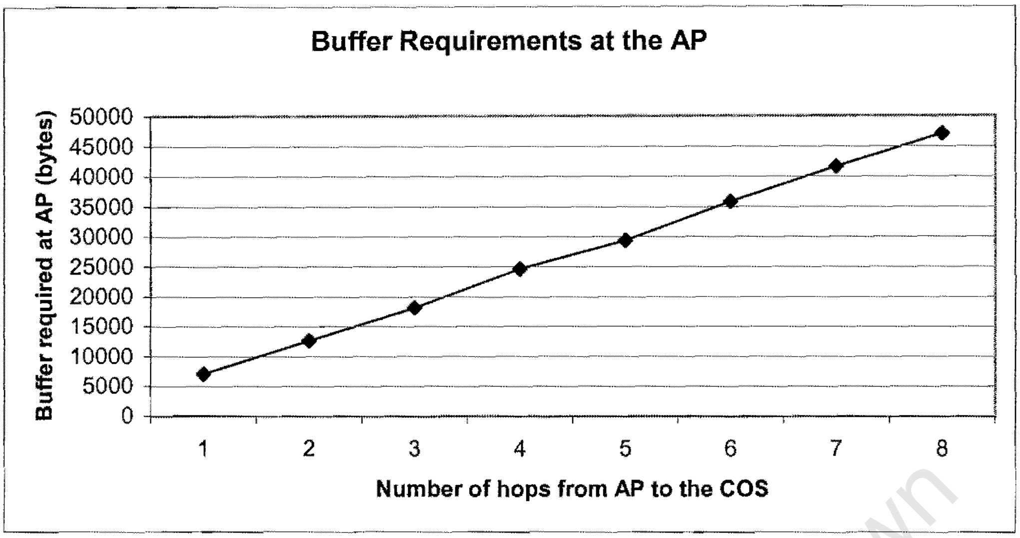


Fig. 6.7 Buffer requirements at the AP for downlink data

6.3.1.4 Buffer requirements at the MT

The buffer requirements at the MT are always the same regardless of the number of hops involved in the handover process. By considering the signaling sequence for the Path Extension scheme, it can be seen that the amount of buffering required at 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. This procedure is the same for both the PR and PE schemes.

The time during which the MT cannot transmit data on the new path is equal to the service disruption time. The buffer requirements at the MT for the Partial Re-Establishment Scheme are given in Fig. 6.8.

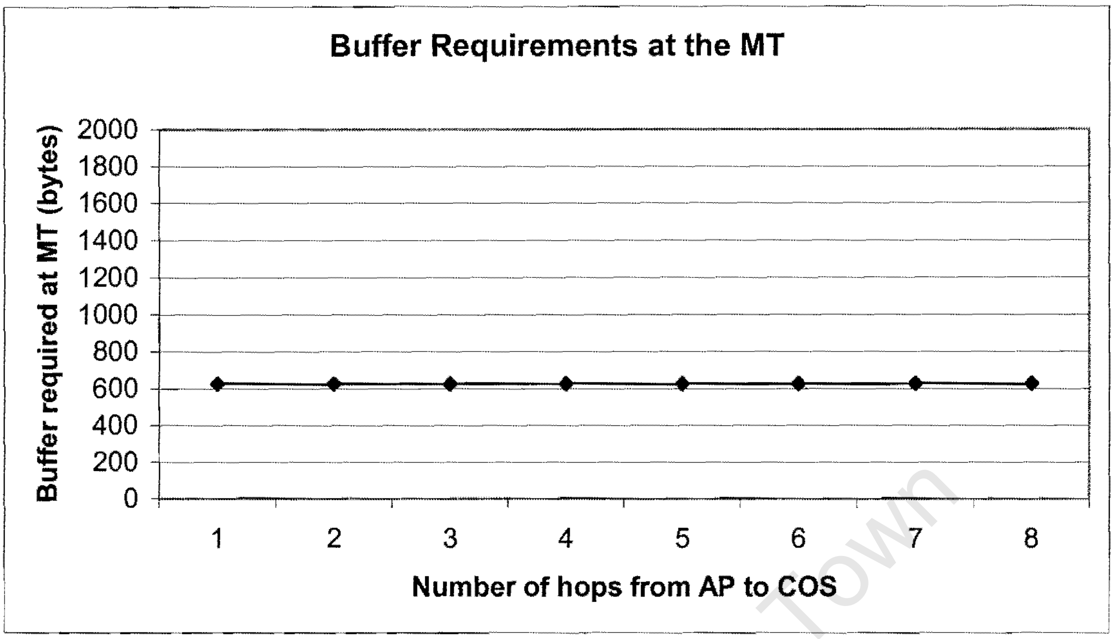


Fig. 6.8 Buffer requirements at the MT for the Partial Re-Establishment Scheme

6.3.2 Performance Evaluation of the Path Extension Scheme

6.3.2.1 Service Disruption Time

The service disruption time for the Path Extension scheme is shown in Fig 6.9. The same procedures were followed to obtain the results as in the Partial Re-Establishment Scheme. Fig. 6.9 shows that the service disruption time for the Path Extension Scheme is constant with the increase in the number of hops between the COS and the AP. The reason for this is that the service disruption time (as in the Partial Re-establishment Scheme) will not be affected by multiple ATM switch handoff, since the disruption is determined only in the last hop between the AP and the MT. From the signaling sequence of the Path Extension Scheme it can be seen that the disruption is only dependent on messages 9 and 10 (Fig. 3.1 and Equation 3.13). Hence, for any number of hops the service disruption time is constant for the Path Extension Scheme. This experimental result confirms the analytical analysis.

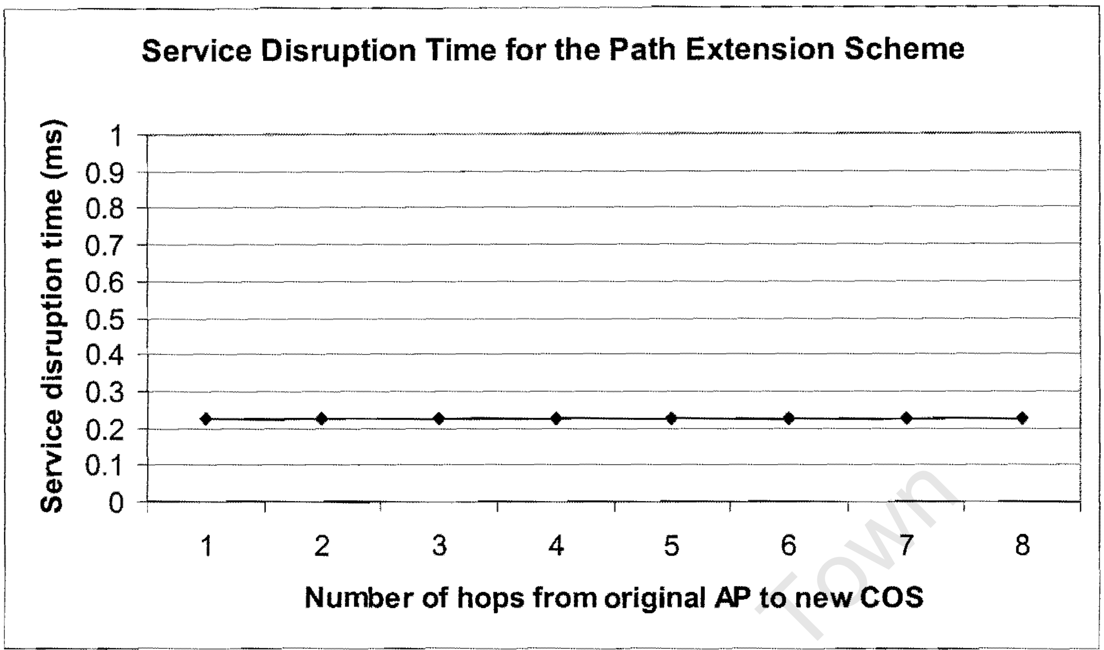


Fig. 6.9 Service Disruption Time for the Path Extension Scheme

6.3.2.2 Handoff Completion Time

The handoff completion time for the Path Extension scheme is shown in Fig. 6.10.

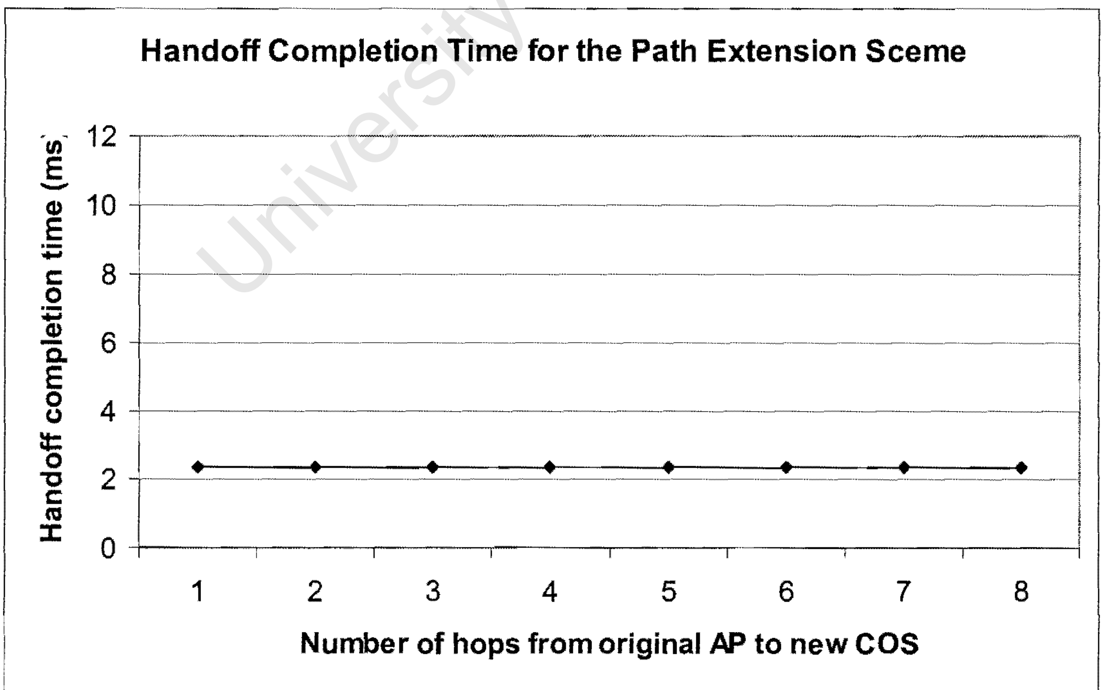


Fig. 6.10 Handoff Completion Time for the Path Extension Scheme

Fig. 6.10 shows that the handoff completion time for the Path Extension Scheme is constant with an increase in the number of hop from the original AP to the COS. The reason for this is that for every handoff, the MT will associate itself with a new AP and only signal to the MES that the new AP is connected to.

6.3.2.3 Buffer requirements at the new AP

The buffer requirements at the new AP for downlink data were also found to be zero. The reason for this is the same as for the Partial Re-Establishment Scheme. In order to determine the buffer requirements at the AP, a delay was added to the handover signaling according to the method described in section 6.3.1.3. Hence, the handover completion time was scaled to the time unit described. The handover completion time is constant, and it follows that the buffer requirements at the AP will also be constant for the Path Extension Scheme. The buffer requirements at the AP for the Path Extension Scheme with the scaled delay are illustrated in Fig. 6.11.

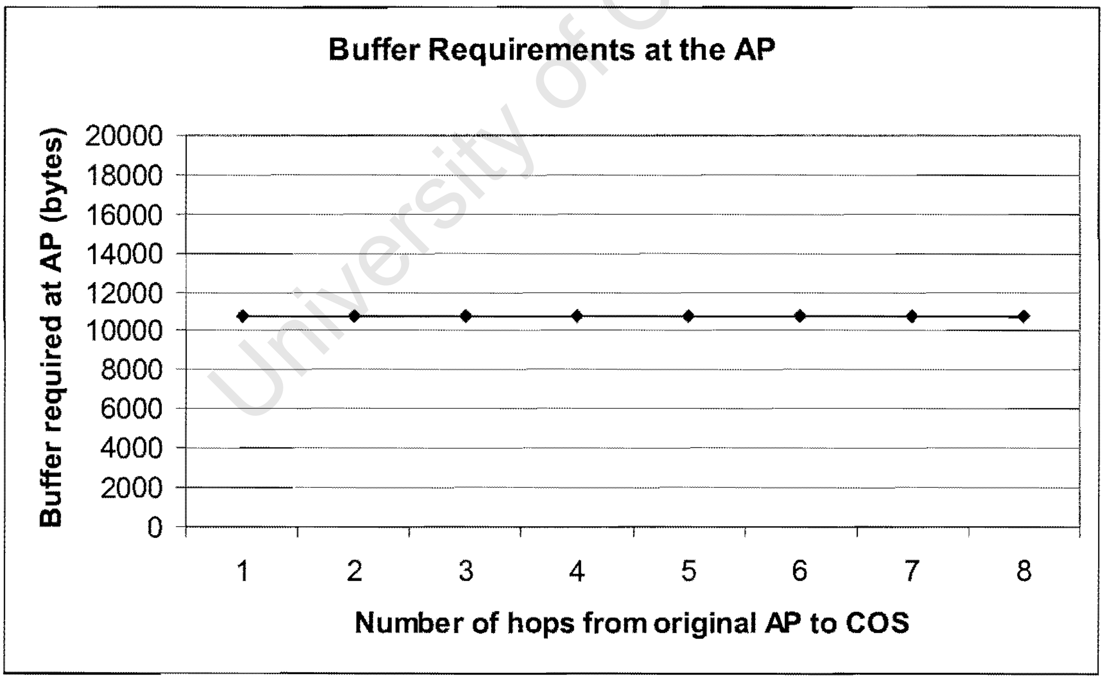


Fig. 6.11 Buffer requirements at the AP for the Path Extension Scheme

6.3.2.4 Buffer requirements at the MT

The buffer requirements at the MT are always the same regardless of the number of hops involved in the handover process. If one looks at the signaling sequence for the Path Extension scheme, it can be seen that the amount of buffering required at 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. This procedure is the same for both the PE and PR schemes.

The time during which the MT cannot transmit data on the new path is equal to the service disruption time. The buffer requirements at the MT for the Path Extension are given in Fig. 6.12.

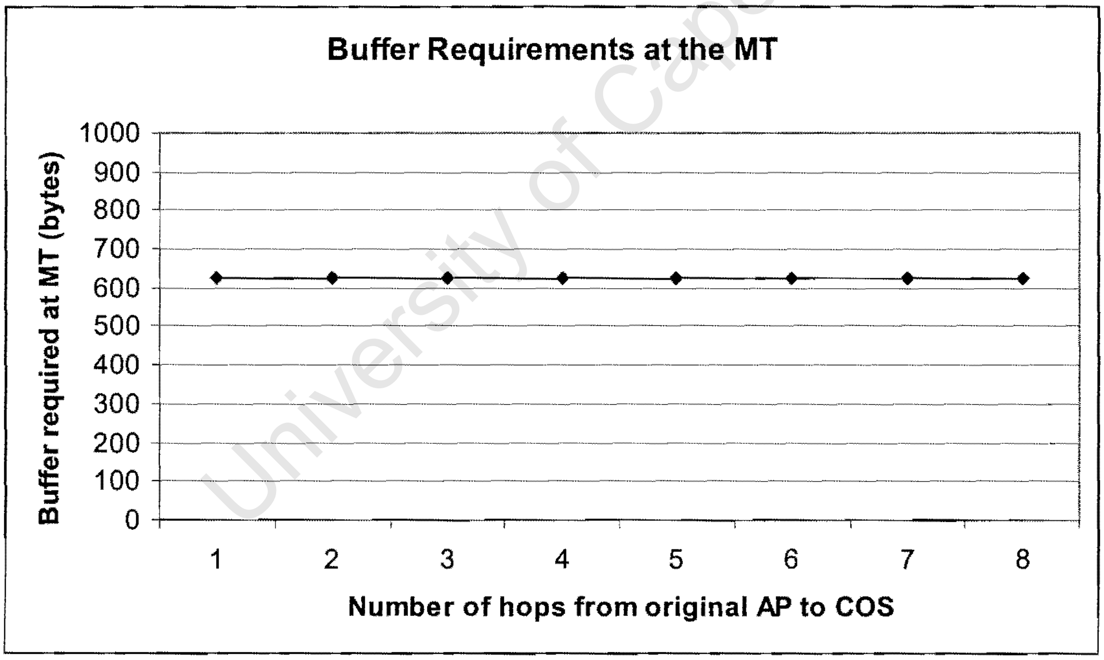


Fig. 6.12 Buffer Requirements at the MT for the Path Extension Scheme

6.4 Comparison of the Partial Re-Establishment and the Path Extension Schemes

Fig. 6.13 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. The reason for this is that the signaling sequences for both schemes are the same as far as the connection to the new AP is concerned. Both service disruption times are constant with an increase in the number of hops between the AP and the COS.

The handoff completion times of the Partial Re-Establishment and Path Extension Schemes are compared in Fig. 6.14. The Path Extension scheme is superior to the Partial Re-Establishment Scheme as far as the handover completion times are concerned. Hence, handover network signaling is significantly reduced for the PE scheme.

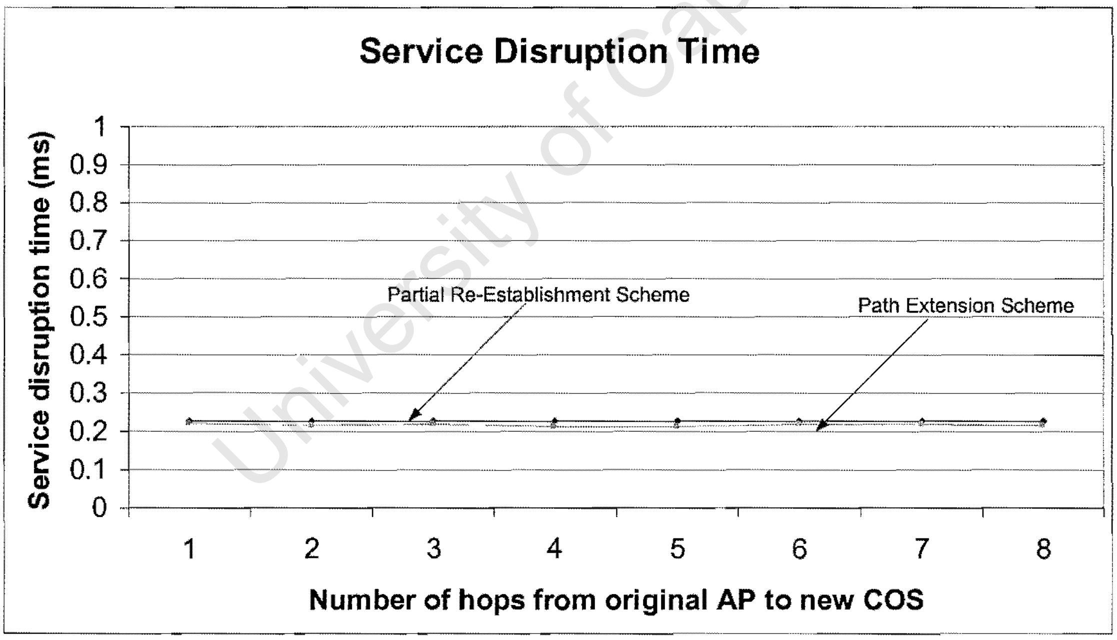


Fig. 6.13 Comparison of the service disruption time

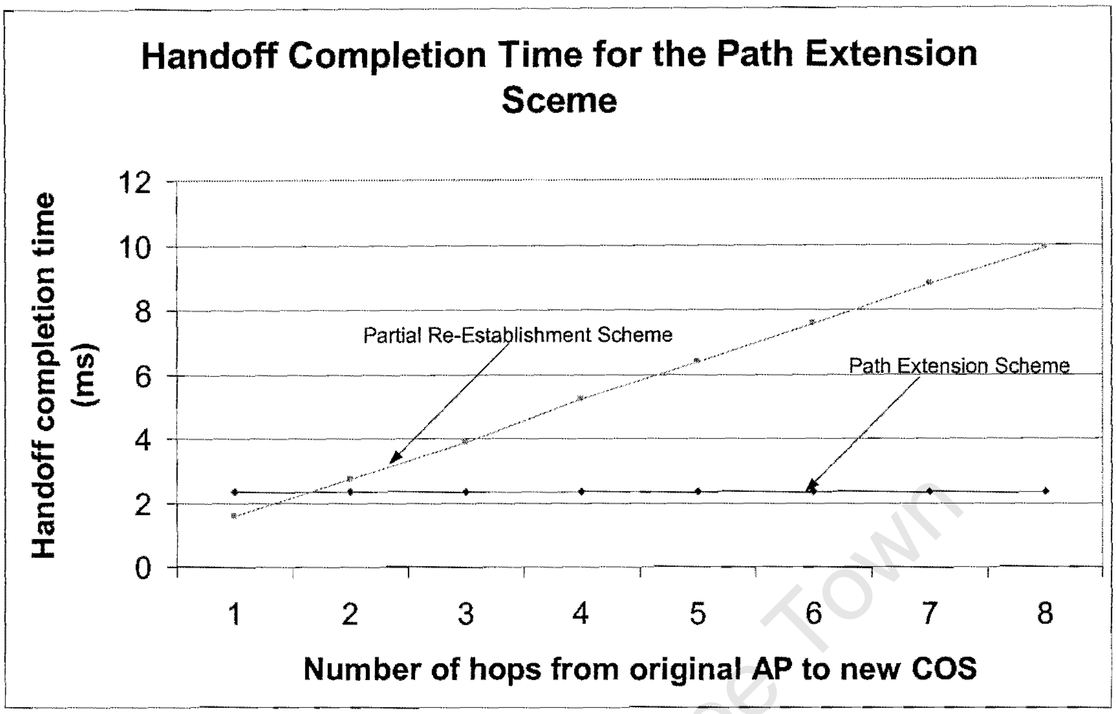


Fig. 6.14 Comparison of the handoff completion time

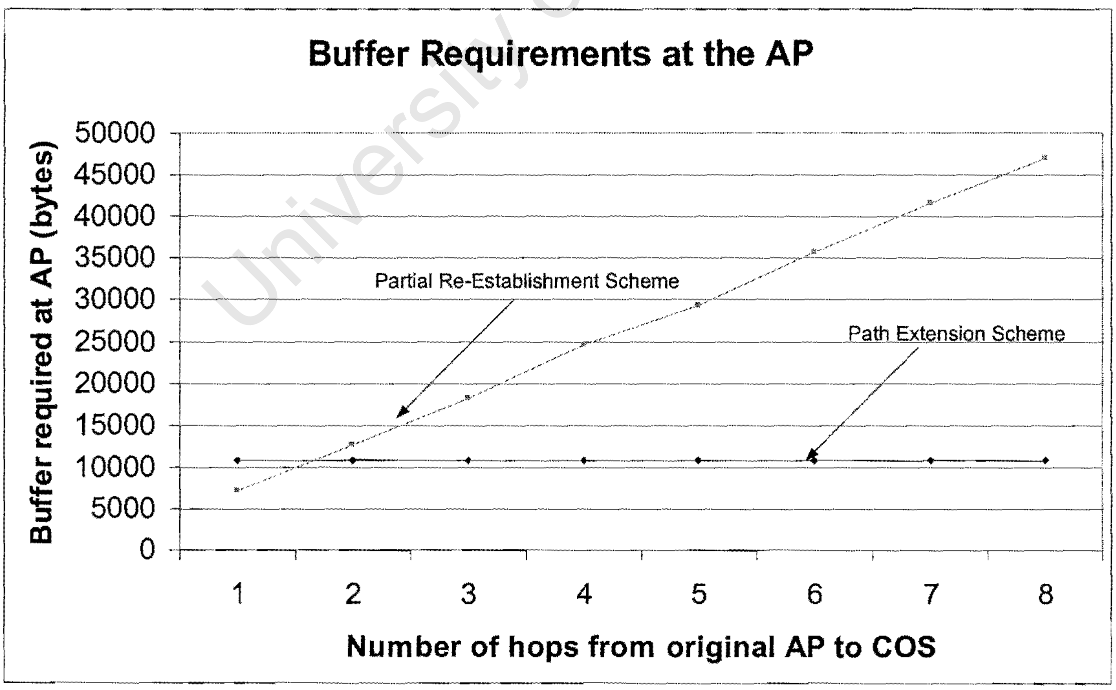


Fig. 6.15 Comparison of the buffer requirements at the AP

Again it can be seen that the buffer requirements at the AP for the Path Extension Scheme is much lower than the Partial Re-Establishment Scheme. Previously it has been hinted that the buffer requirements at the AP is directly related to the service disruption time. This relationship can also be seen by comparing Fig. 6.14 and 6.15.

The buffer requirements at the MT for both the Partial Re-Establishment and Path Extension Schemes are identical.

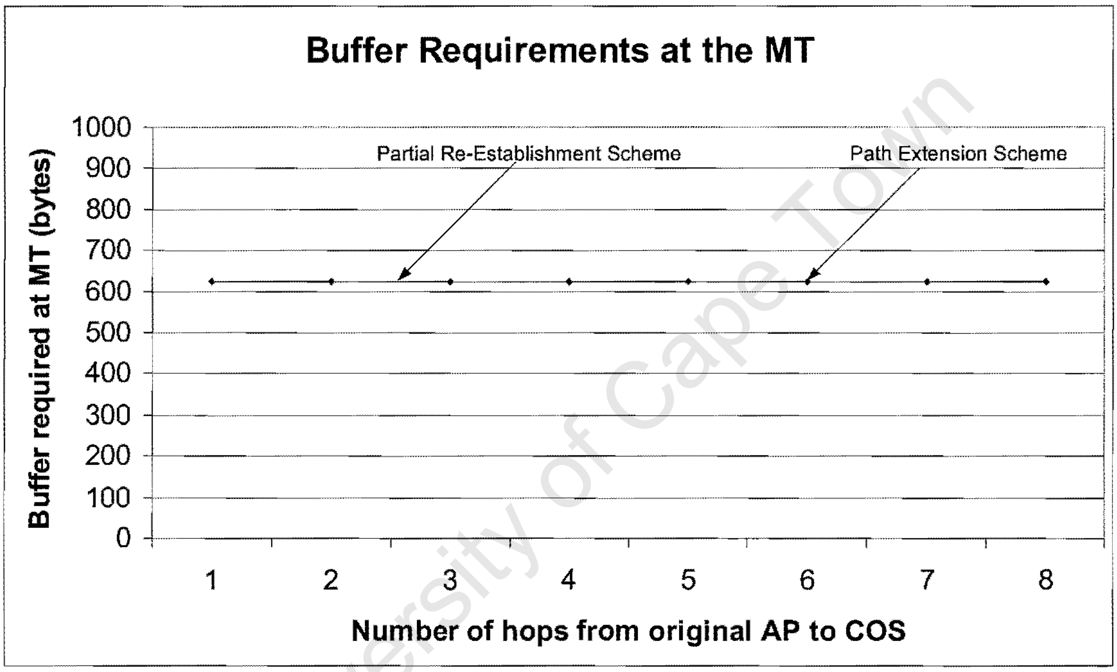


Fig. 6.16 Comparison of the buffer requirements at the MT

6.5 COS position variation for the PR Scheme

The same procedure as in section 3.5 was followed experimentally by varying the position of the COS, i.e. the position of the COS was not centrally located between the new and the old AP. Hence we considered the following scenario:

We have x hops between the new AP and the old AP, with the COS y hops from the new AP and ($x - y$, $x > y$) hops from the old AP. Thus, we performed handover operations by varying the position of the COS between the new and the old AP, while keeping the total number of hops between the old and new AP constant. The handover completion time for this scenario is illustrated in Fig. 6.17.

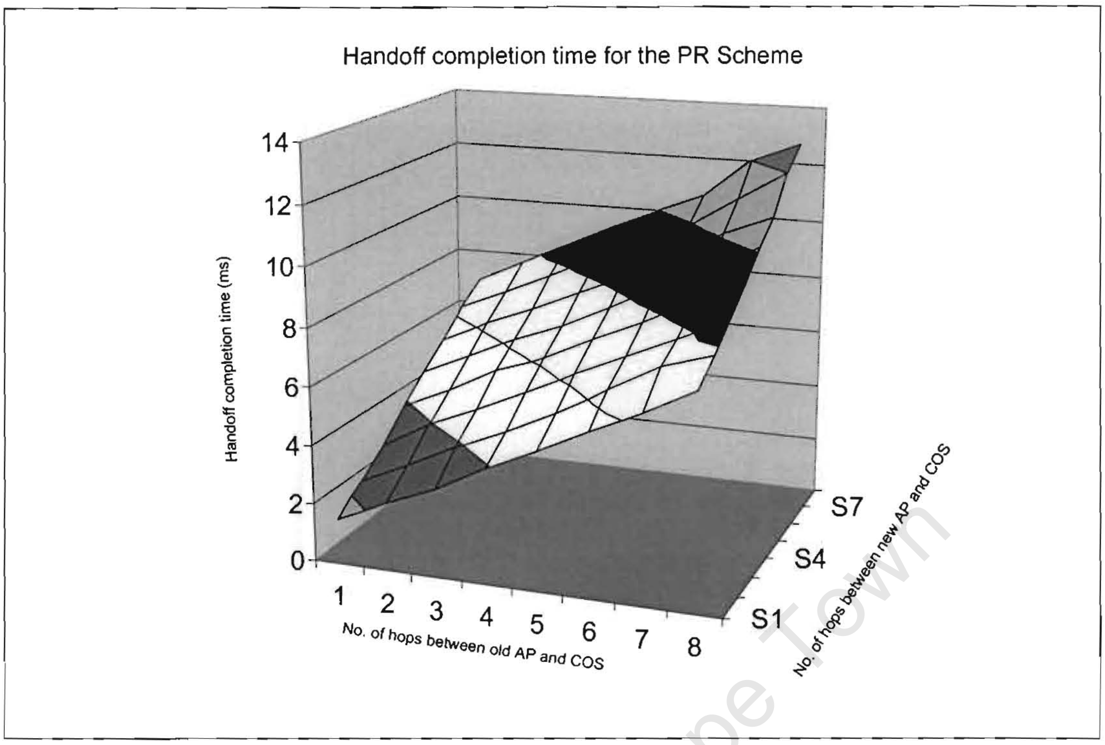


Fig. 6.17 Handoff completion time for the PR scheme with COS location varied

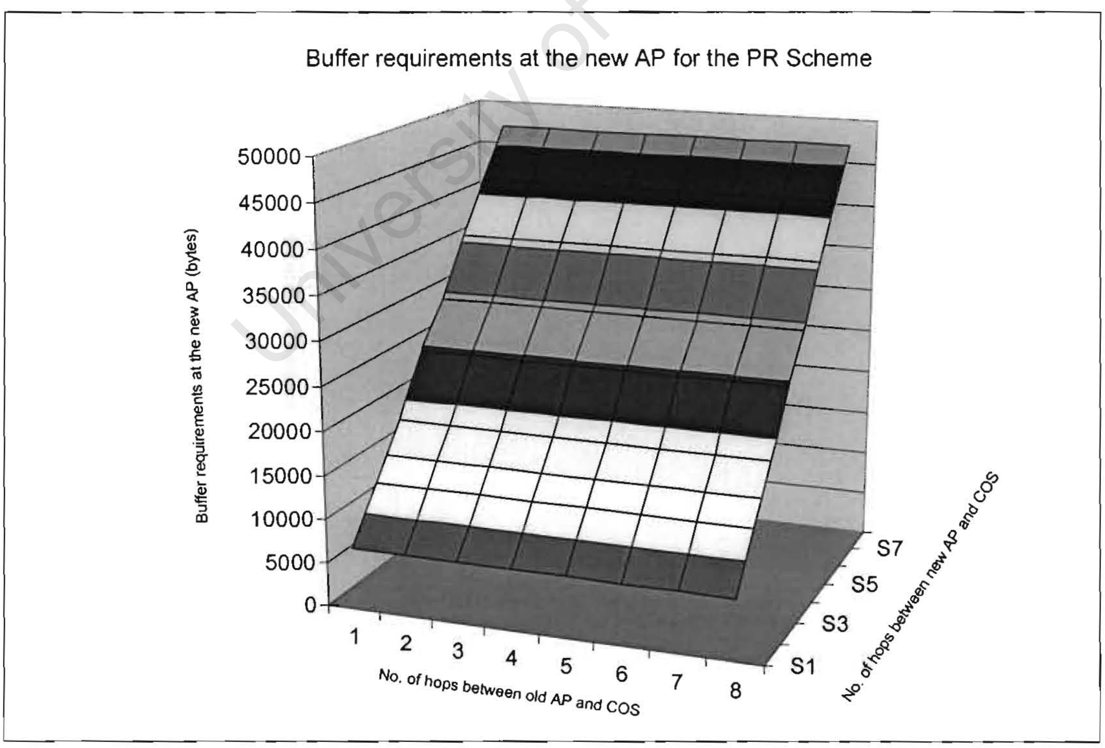


Fig. 6.18 Buffer requirements at the new AP with the COS location varied

As with the analysis carried out in chapter 3, the buffer requirements at the new AP for the downlink data is only dependent on the number of hops between the COS and the new AP, and the number of hops between the old AP and the COS has no effect on it. By increasing the number of hops between the old AP and the COS, we only increase the handover completion time. As previously mentioned in chapter 3, by comparing Fig. 6.17 and 6.18, one can conclude that it is better to allocate the position of the COS as close to the new AP as possible in order to enhance the overall performance of the PR scheme. Although it was not the initial intention to measure this result, the author made this observation by a careful consideration of how the handoff completion time affects the buffer requirements at the new AP. This result could be used in COS location algorithms in order to enhance the performance of the PR scheme.

Chapter 7

Conclusions and Recommendations

7.1 Conclusions

This research focused on designing and implementing a functional WATM architecture and to evaluate the performance of handoff schemes in WATM networks. The preceding chapters have presented the design and implementation of the functional WATM architecture and signaling framework to support handover in the WATM network. Conclusions made during this research are presented in this section.

The handover signaling protocols studied in this project are based on the PR and PE schemes. The protocols were studied individually in order to quantify their characteristics. However, it was shown why the two schemes could be used as a combination to facilitate the Two Phase Handoff scheme.

A quantitative comparison of the PE and PR schemes were presented using analytically derived formulas for the following performance metrics: (i) Service disruption time, (ii) Handoff completion time, (iii) Buffer requirements at MT, and (iv) Buffer requirements at AP required during a handoff.

The analysis provided a description of what the effect of different network topologies would have on the performance metrics of the handoff schemes. Of particular interest was to determine how the performance metrics would be affected with an increase in the number of hops between the MES and COS. The analysis implied that the handoff completion time of the PE scheme is not affected by multiple switch handoffs. However, for the PR scheme, it increases with an increase in the number of hops between the MES and the COS. Although the PE scheme has a fast and consistent handoff completion time compared to the PR scheme, it sacrifices an optimal path for an increase in handoff speed, whereas the PR scheme sacrifices a fast handoff completion time for an optimal path. The buffer requirements at the MT for both schemes are constant with an increase in the number of hops between the COS and MES, however, the buffer requirements at the new AP is constant only for the PE scheme and increases linearly for the PR scheme. The Two Phase handoff scheme uses the PE scheme as a first phase to facilitate a fast handoff and the PR scheme in the second phase to provide an optimal path.

The analysis also showed that if the number of hops between the old MES and the new MES are kept constant, and only the position of the COS is varied for the PR scheme, that the handover completion time remains constant. However, this had a striking effect on the buffer requirements at the new AP. The analysis showed that the buffer requirements at the new AP increases linearly with an increase in the number of hops between the COS and the new MES only, and that by increasing the number of hops between the COS and the old MES had no effect on it. Although the project does not concern itself with COS discovery algorithms, this results could be used to enhance the performance of the PR scheme by incorporating it into COS discovery algorithms.

The project included an experimental implementation of the WATM network. This required the building of a prototype WATM network with existing ATM switches and implementing the handover signaling entities at both the access and network sides. This involved developing a control-plane signaling entity at the various WATM functional entities to provide the necessary handover signaling functionality. The signaling entity was integrated with the WATM user-plane functional components.

It was shown how a fixed ATM switch with minor modification, can be changed into a MES that is capable to support a fast handoff at the ATM level.

It was critical for future work that the signaling entity framework developed be based on a modular design methodology ensuring that future research employ this framework with minimal modification effort. This was achieved by selecting a modular design approach in developing the control plane handover signaling framework.

Having developed the handover signaling entities, it was necessary to design an evaluation platform to verify the performance of the system. After successful implementation of the handover signaling entity framework, the system was tested to determine the functional correctness within the developed system.

To evaluate the functional correctness of the WATM implementation, several tests were conducted. In these tests, streams of packets were transmitted through the switch from one end systems and received at another. Packets were logged at both end stations in order to assess correctness of delivery when a handoff occurred. By examining these log files, it was concluded that the MES, APs and MTs were functioning correctly.

A major obstacle was experienced with the Linux and NetBSD operating systems as a platform for the development of the software components. These operating systems could only support a maximum time resolution of 10ms. Even if ample processing capability is available, it is not possible for a module operation to complete in under 10ms. This required that the results be calibrated by normalizing the time-scale to time units, where each time unit represented 10ms.

An extensive investigation was done to assess the performance parameters of the PE and the PR Schemes on the evaluation platform. As with the analysis, the service disruption time, time taken to complete a handoff and extra buffering needed to avoid data loss due to rerouting were important performance issues that were considered as they affect the quality of service of the connections.

The conclusions drawn from the experimental results were that they are consistent with the analysis. However, it was found that with the evaluation platform the buffer requirements at the new AP were always zero. The reason for this was due to the maximum timer resolution imposed by the operating system.

7.2 Recommendations

Based on the research performed during this thesis, the following recommendations can be made:

In the introduction of this thesis it was stressed that handoff must be supported with low cell loss, latency and control overhead. The QoS guarantees must be maintained for each connection during and after MT migration, i.e. dynamic resource allocation, QoS provisioning and handoff control. Also, a MT can experience varying radio and network environments during migration, thus necessitating QoS renegotiation when the existing guarantees can no longer be met. This study assumed that the QoS of the connection is not affected during a handover and that it remained unchanged during and after handover. This is not always the case and needs to be investigated further.

The current design developed in this project only supports one handover instance at any point in time. If more than one handover are requested simultaneously by multiple MTs connected to the same AP, each individual handover procedure would follow the same sequence of events. However, an additional module would have to be incorporated at the AP in order to provide a scheduling mechanism for simultaneous handoff requests. An AP scheduling architecture would have to support and guarantee the QoS of connections from various MTs. The evaluation platform could be extended to include support for simultaneous MT handoff requests. Thus, attention must be paid to the scheduling of the simultaneous handoff request at the AP during handover. An area of further research is to investigate the properties of AP scheduling architectures to support the QoS of various MT connections during simultaneous handovers.

Although not directly related to the work performed in this project, it should also be a requirement of the WATM network to support the TCP/IP protocol over WATM networks because of the widespread use of the Internet (based on the TCP/IP protocol). An extension of this project is to study the effect that the two handoff protocols will have on TCP/IP performance during a handover.

Another issue regarding TCP/IP over ATM is related to the radio. The question in this case is where the segmentation and reassembly for ATM will be made, or where the ATM adaptation layer will be located. If this location is the AP, requirements on the radio can be simplified since then packet sizes can be larger (IP packets are large). The main disadvantage of this approach is a more complicated AP whose complexity increases with the average number of active users. A need therefore exist to investigate the possibility to provide an infrastructure for ATM cell transport over wireless channels.

During the course of this study it was mentioned that the buffer requirements at the new AP for the downlink data is only dependent on the number of hops between the COS and the new AP, and that the number of hops between the old AP and the COS has no effect on it. By increasing the number of hops between the old AP and the COS, we only increase the handover completion time. Therefore, one can conclude that it is better to allocate the position of the COS as close to the new AP as possible in order to enhance the overall performance of the PR scheme. This result could be used in COS location algorithms in order to enhance the performance of the PR scheme and therefore the optimization process in the second phase of the Two Phase Handoff Scheme. Further investigation needs to be undertaken to incorporate this finding with existing COS discovery algorithms.

References

1. ATM Forum Technical Committee, "Draft Wireless ATM Capability Set 1 Specification", ATM Forum BTD-WATM-01.13, 2000.
2. G.A. Awater and J. Kruys, Wireless ATM – an overview. Mobile Networks and Applications 1 (1996) 235 – 243.
3. IEEE Journal on Selected Areas in Communications, Special Issue on Wireless ATM, 1997.
4. A. Acompora and M. Naghshineh, An Architecture and methodology for mobile-executed handoff in cellular ATM networks, IEEE Journal on Selected Areas in Communications 12(8) (October 1994) 1365 – 1375.
5. R. Yuan, S.K. Biswas and D. Raychaudhuri, Signaling and Control Architecture for Mobility Support in Wireless ATM Networks.
6. P. Marichamy, S. Chakrabarti and S.L Maskara. Performance Issues in Rerouting for Handoff in Wireless ATM, 2000.
7. R. Ramjee, T. La Porta, J. Kurose, D. Towsley. Performance Evaluation of Connection Rerouting Schemes for ATM-based Wireless Networks. IEEE/ACM Transactions on Networking, 1998.
8. K. Keeton, B. Mah, S. Seshan, R. Katz and D. Ferari. Providing Connection-Oriented Network Services to Mobile Hosts. Proceedings of USENIX Symposium, August 1993.
9. H. Mitts, H. Hansen, J. Jimmonen, S. Veikkolainen, Lossless Handover for WATM, Proceedings of MobiCom 1996.
10. D.K. Kim and C.K. Toh, Mobile multicasting in wireless ATM networks. Mobile Networks and Applications 5 (2000) 103 – 110.
11. S.J. Lee and D.K. Sung, A New Fast Handoff Management Scheme in ATM-based Wireless Mobile Networks. IEEE (1996) 1136 – 1140.
12. A. Hac and J. Peng, A Two-phase combined QoS-based handoff scheme in a wireless ATM network. International Journal of Network Management (2001) 309 – 330.

13. J. Elbergali and N. Ventura, Performance evaluation of a new fast path re-routing for inter-cluster handoff scheme in Wireless ATM networks, IEEE ICCS (2002).
14. A. Alles, The Next-Generation ATM switch : From Testbeds to Production Networks, Cisco Systems Inc, document number 119WP (1995).
15. H. Kim, H.J. Chao, Design of an ATM switch for handoff support. Wireless Networks 6 (2000) 411 – 419.
16. M. Veeraraghavan, T.F. La Porta, R. Ramjee. A Distributed Control Strategy for Wireless ATM Networks, in Wireless Networks Journal, October 1995.
17. R. Cohen B. Patel, A. Segall, Handover in micro-cell packet switched mobile networks, Wireless Networks 1 (1996).
18. M.J. Karol, K.Y. Eng and M. Veeraraghavan, BAHAMA: A broadband ad-hoc wireless ATM local area network, in: *Proc. ICC '95* (June 1995) pp. 1216–1223.
19. B.A. Akyol and D.C. Cox, Rerouting for handoff in a wireless ATM network, IEEE Personal Communications (October 1996) 26–33.
20. C.-K. Toh, Crossover switch discovery for wireless ATM LANs,ACM Journal on Mobile Networks and Applications 1(2) (October1996) 145–165.
21. C. Chrysostomou, A. Pitsillides, F. Pavlidou, A Survey of Wireless ATM Handover Issues, International Symposium of 3G Infrastructure and Services - 3GIS Athens, Greece. 2001.
22. A. Acharya, J. Li, B. Rajagopalan, D. Raychaudhuri, "Mobility Management in Wireless ATM Networks", IEEE Communications Magazine, vol. 35, no. 11, 1997, pp. 100-109.
23. H.-J. Vogel, "Robust and Soft: Handover Design for High-tier Mobile ATM systems", in Proceedings of Wireless'99, 1999.
24. C.-K. Toh, "A Hybrid Handover Protocol for Local Area Wireless ATM Networks", MONET – Mobile Networks and Applications, vol. 1, no.3, 1996, pp. 313-334
25. H.-J. Vogel, "Handover Switching in Mobile ATM Networks", in Proceedings of European Personal Mobile Communications Conference EPMCC'97, 1997.

26. H.-J. Vogel, "Virtual Connection Handover in ATM-based Cellular Networks", in Proceedings of European Conference on Networks and Optical Communications NOC'98, 1998, pp. 298-305.
27. H.-J. Vogel, "A Networking Concept for Wide Area Mobility and Fast Handover in ATM Networks", in Proceedings of IEEE Globecom'97, 1997, pp. 1124-1128.
28. H. Mitts, H. Hansen, J. Immonen, S. Veikkolainen, "Lossless Handover for Wireless ATM", MONET – Mobile Networks and Applications, vol. 1, no. 3, 1996, pp. 299-312.
29. A. Kaloxylos, G Papageorgiou, P. Papageorge, L. Merakos, "Smart Buffering Technique for Losless Hard Handover in Wireless ATM Networks", ACM/Baltzer Mobile Networks and Applications, 1(3), December 1996
30. O.T.W. Yu, V.C.M. Leung, "Connection Architecture to support efficient handoffs over an ATM/B-ISDN PCN", ACM Journal on Mobile Networks and Applications, 1996, pp. 457 -473.
31. Y.B. Lin, A.C. Pang, " Comparing Soft and Hard Handoffs", IEEE Trans. On Veh. Technology.
32. F. Jiang, T. Kakola, "Wireless ATM : Handover Issues", Proceedings of the Personal, Indoor and Mobile Radio Communications. IEEE 1998.
33. F. Jiang, T. Kakola, "Microcellular Handover in WATM", Proceedings of the 32nd Hawaii International Conference on Systems Science, 1999.
34. J.R. Elbergali, "A Framework for Fast Handoff Schemes in Wireless ATM Networks", PhD. Thesis, University of Cape Town, 2002.
35. K.H. Salah, "A Novel Handoff Scheme for Wireless ATM Networks", PhD. Thesis, Illinois Institute of Technology, 2000.
36. WUGS – Washington University Gigabit Switch, URL: http://www.arl.wustl.edu/gigabitkits/kits.html
37. D. Wu, J.D. De Hart, The Gigabit Network Switch Controller, Washington University ARL, June 1998.
38. J.S. Turner, A Gigabit Local ATM Testbed for Multimedia Application, Washington University, ARL Working Note ARL-94-12, Version 1.2, June 1996.

39. O. Beal, J.D. De Hart, Jammer Language Discription, Washington University ARL, June 1998.
40. D. Wu, J.D. De Hart, Node Controller Managed Object (NCMO) and Node Controller Communications Protocol (NCCP), Washington University ARL, June 1996.
41. Waël A. Hassan, "Report on Handoff and Mobility Management in Wireless Networks", M.CS. University of Ottawa.
42. N. Narang, S. Kasera, "ATM SVC Signaling Stack : Implementation Guidelines"
43. ITU-T Recommendation Q.2110, B-ISDN Signaling ATM Adaptation Layer overview description. 1994.
44. Fore Systems Inc., "ATM Networks", Technical Report, 1992.
45. Internet Protocol – DARPA IP Specification, RFC 793, 1981.
46. Transmission Control Protocol – DARPA TCP Specification, RFC 791, 1981.
47. Netcomm Ltd.,"DV 2 – Switch Installation Managers Guide, V 2.0", 1993.
48. B. Stiller, "A Survey of UNI Signaling Systems and Protocols for ATM Networks"
49. ATM-Forum 94-0998, "Signaling Subgroup Meeting Notes, UNI 4.0 Features, 1994.
50. ITU-T Recommendation I.361, B-ISDN ATM Layer specification. 1994
51. UNI 3.1 Specification, ATM Forum, September, 1994.
52. L. Frelechoux, "PNNI Addendum for Network Call Correlation Identifier", ATM Forum, BTD-CS-NCCI-01.04, September, 1999.
53. Traffic Management Specification, Version 4.0, AF-TM-0056.00, April 1996, ATM Forum.
54. "Network Timing Protocol", http:/www.eecis.udel.edu/~ntp/ntp_spool/html, May 1998.
55. N. Passas and L. Merakos, "Wireless ATM", The Wiley Encyclopedia of Tellecomunications.
56. J. B. Cain and D. N. McGregor, "A Recommended Error Control Architecture for ATM Networks with Wireless Links", *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 16-28, Jan. 1997.

57. N. Guo and S. D. Morgera, "Frequency-Hopped ARQ for Wireless Data Services", *IEEE Journal on Selected Areas in Communications,* vol. 12, pp. 1324 -1337, Oct. 1994.

58. H. Liu, H. Ma, M. El Zarki a and S. Gupta, "Error control schemes for networks: An overview" Mobile Networks and Applications 2 (1997) 167–182.

Appendix A

The Accompanying CD-ROM

A CD-ROM accompanies this thesis. This CD-ROM contains the following information:

- This thesis document in ".doc" format can be found in the "Thesis Document" directory.
- Many of the documents referenced in this thesis can be found in the "Referenced papers and other useful documents" directory.
- Source code for the Handover signaling entities at the various network elements can be found in the "Handover framework source" directory.
- The images used in this thesis have been included in various formats in the "Thesis Document\Images" directory.

Appendix B

B.1 ATM-on-Linux

This section presents a guide to the main steps taken during this study to learn about ATM-on-Linux and to construct the MES and network signaling.

The operating systems used in this study are as follows for the various WATM network elements:

1) Linux kernel version 2.4.18 for the fixed host, MES and MT.
2) NetBSD 1.4.1 for the AP's.

B.2 Installation of Linux

The basic Linux operating system can either be installed from a FTP site or a distribution CD. Refer to the 'Linux Installation and Getting Started' book or the University of Cape Town LEG (Linux Enthusiasts Group) site http://www.leg.uct.ac.za/starting_out.html for

information to get started.

Install Redhat 8.1 with linux kernel version 2.4.18. Information and help for installation can be obtained from the UCT LEG website.

B.3 Add ATM-on-Linux support

In order to install ATM support on Linux, you need the following components:

• The package itself linux-atm-2.4.0.tar.gz (can be downloaded from the ATM-on-Linux main webpage).

• Linux kernel version 2.4.18 (which you are now hopefully running)

For installation of the ATM-on-Linux software, see ATM on Linux users guide Release 0.59.

Appendix C

```
#Manipulate bidirectional connections through wugs

#initialize port

proc initialize ()
        int port
        int numports
        port = 0
        numports = 8
                while ($port < $numports)
                write mr $port 2 0 128 32 0 256 1 1 1 1 100000 0
        port ++
        done
end

# Set port 1A->0A and 1B->0A
# and port 0A->1A or 0A->1B or 0A->0A based on VPI/VCI
# NOTE: OC-3 duplex card does not allow TX/RV on same VCI
# OBSCURE POINT: To access upper half of OC-3 duplex (B port),
# the line card uses bit VPI[7] (ie VPI >128 = B port).

proc AddPvc(int portNum1, int vpi1, int vci1, int portNum2, int vpi2, int vci2)
write vcxt $portNum1 $vci1 1 2 1 0 0 1 0 0 0 0 0 0 $vpi2 $vci2 0 0 0 0 $portNum2 0
write vpxt $portNum1 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
write mr $portNum1 2 $portNum2 128 32 0 256 1 1 1 1 100000 0
end

#
# Add the following PVC
# portNum1/vpi1/vci1    <==>  portNum2/vpi2/vci2
# Assumes that it is a bi-directional connection.
#
proc AddPvc2(int portNum1, int vpi1, int vci1, int portNum2, int vpi2, int vci2)
     write mr $portNum1 2 $portNum2 128 32 0 256 1 1 1 1 100000 0
     write mr $portNum2 2 $portNum1 128 32 0 256 1 1 1 1 100000 0
     write vpxt $portNum1 $vpi1 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
     write vcxt $portNum1 $vci1 1 2 1 0 0 1 0 0 0 0 0 0 $vpi2 $vci2 0 0 0 0 $portNum2 0
     write vpxt $portNum2 $vpi2 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
     write vcxt $portNum2 $vci2 1 2 1 0 0 1 0 0 0 0 0 0 $vpi1 $vci1 0 0 0 0 $portNum1 0
end
proc RmPvc2 (int portNum1, int ppi1, int vci1, int portNum2, int vpi2, int vci2)
        write mr $portNum1 2 0 128 32 0 256 1 1 1 1 100000 0
```

```
        write mr $portNum2 2 0 128 32 0 256 1 1 1 1 100000 0
        write vcxt $portNum1 $vci1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
      write vcxt $portNum1 $vci1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

end
proc Recycle(int portNum1, int vpi1, int vci1, int portNum2, int vpi2, int vci2)
write mr $portNum1 2 $portNum1 128 32 0 255 1 1 1 1 1048576 0
write vpxt $portNum1 $vpi1 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
write vcxt $portNum1 $vci1 1 2 1 1 0 0 0 0 0 0 0 0 $vpi2 $vci2 0 0 0 0 $portNum2 0
end
initialize()
#include AddPvc.js
#rm()
include control_vcs_0hop.js
include control_vcs_1hop.js
include control_vcs_2hop.js
include control_vcs_3hop.js
include control_vcs_4hop.js
include control_vcs_5hop.js
include control_vcs_6hop.js
include control_vcs_7hop.js
include control_vcs_PE.js
#include data_vcs_0hop.js
include data_vcs_1hop.js
include data_vcs_2hop.js
include data_vcs_3hop.js
include data_vcs_4hop.js
```

Appendix D

Handover Signaling Entities at the various network elements

This section describes the message flow diagrams of the signaling entities (presented as flow diagrams) at the various network elements for both the PR and PE schemes.

D.1 Handover Protocol Flow Diagrams for the PE Scheme

Having introduced the various network signaling and data elements, we now have a closer look at the handover signaling processing within these entities for the PE scheme presented in the previous chapter.

The flow diagrams provide more details and insight into handling implementation issues and addressing different scenarios. The flow diagrams at the old MES and the new MES are different. They are also different for the old AP and the new AP. We therefore consider these flow diagrams separately. Also note that the flow diagram at the old MES will handle both intra- and inter-switch handoffs.

D.1.1 Message flow within the Mobile Terminal Signaling Entity

The MTSE is implemented as a software entity on the MT. The flow diagram for the handover signaling stack at the MT is illustrated in Fig. D1.

The MT signaling stack has an extra module called the Radio Link Monitor that is used for radio link monitoring. When the Radio Link Monitor detects that the radio receive level drops below a set threshold value, it triggers a handover request to the new AP.

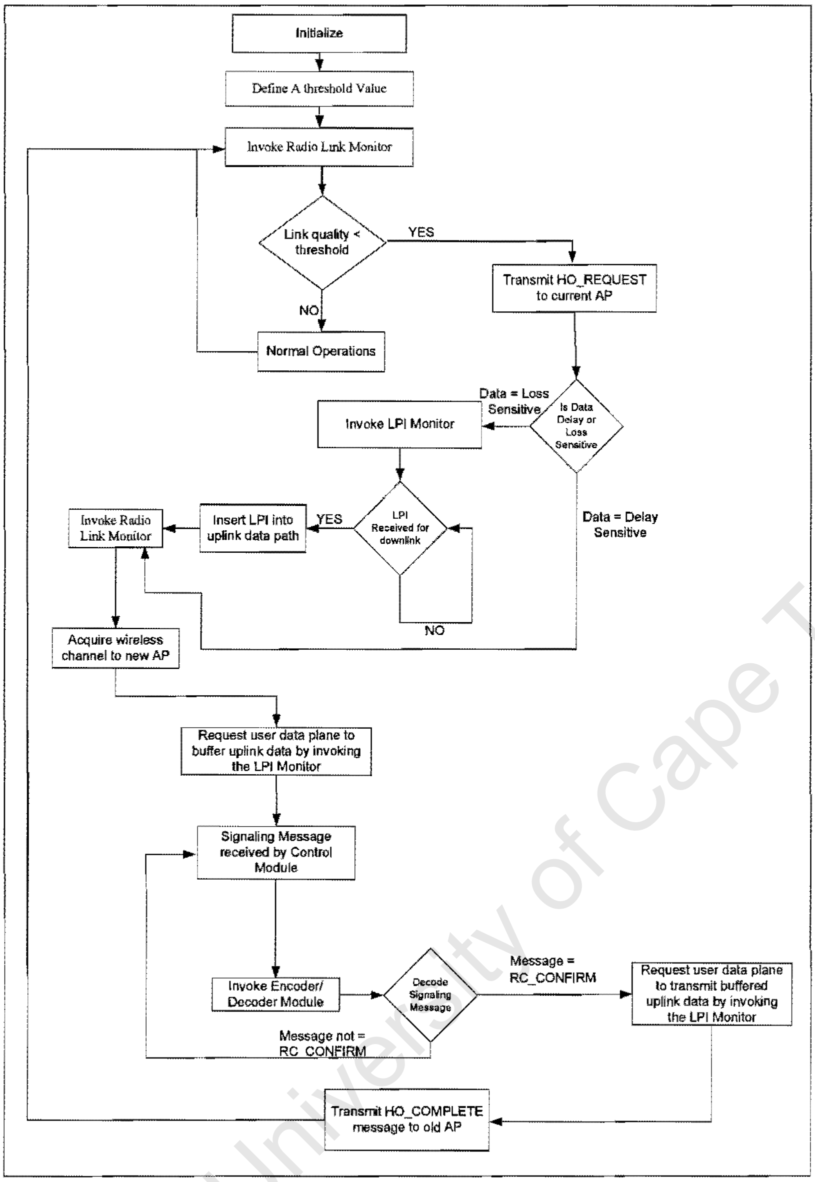


Fig. D1 MTSE flow diagram

D.1.2 Message flow within the Access Point Signaling Entity

The APSE is implemented as a software entity on the old and new APs. The flow diagram for the handover protocol at the old AP is illustrated in Fig. D2.

The old AP receives a handoff request from the MT and informs the MES that a handoff is requested. The old AP then waits for the HO_CONFIRM message from the MES, transmits it to the MT and waits for confirmation from the MT that the handoff is complete and that the channel to the MT can be released. The APSE at the old AP ends by transmitting a handoff completion and connection release message to the MES.

The flow diagram for the handover protocol at the new AP is illustrated in Fig. D3. The new AP receives a handoff channel establishment request from the COS via the new MES and allocates the necessary resources. It then confirms the resource allocation by sending a message to the COS via the new MES and implements a buffer that will intercept all downlink data for the connection until it receives a radio channel establishment request from the MT. After sending a radio channel confirmation message to the MT, it transmits the buffered downlink data and all newly arriving downlink data to the MT.

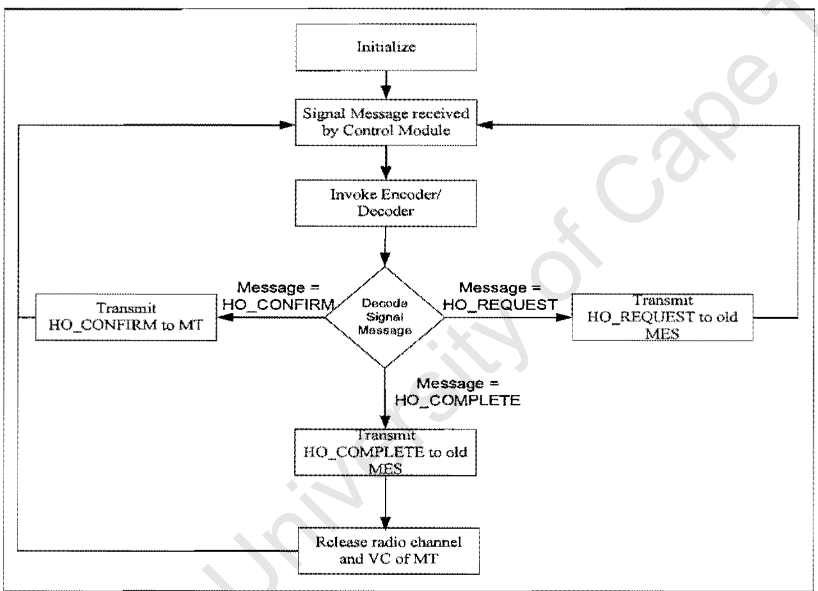


Fig. D2 APSE handover flow diagram at the old AP

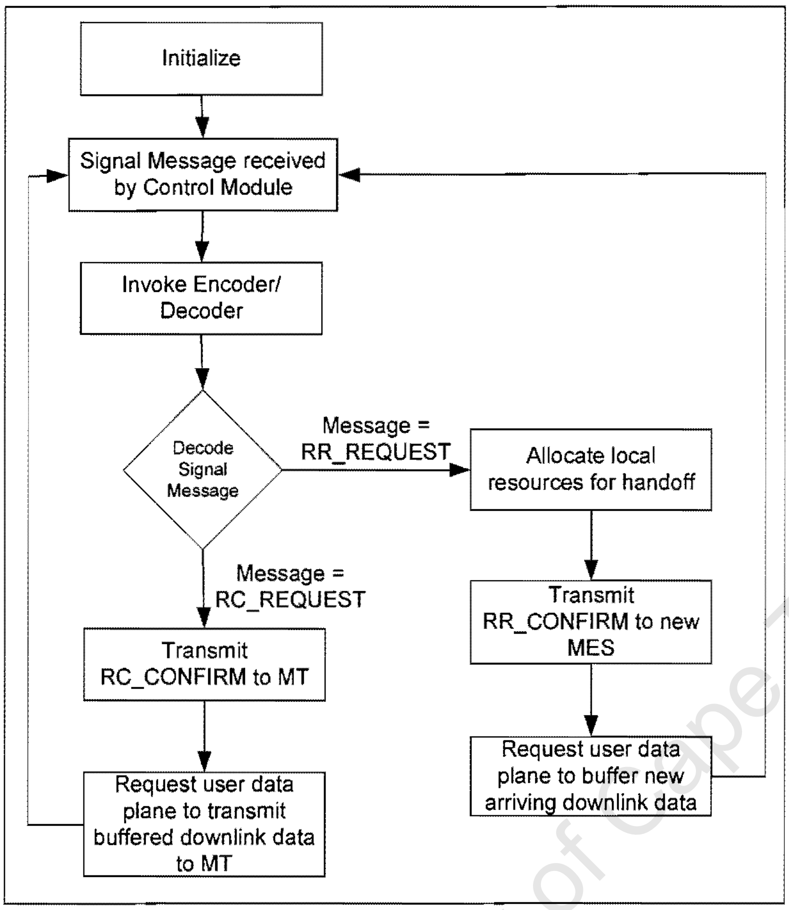


Fig. D3 APSE handover flow diagram at the new AP

D.1.3 Message flow within MES Signaling Entity

The MESSE is implemented as a software entity on the old and new MES. The flow diagram for the handover protocol at the old and new MES is illustrated in Fig. D4 and Fig. D5 respectively.

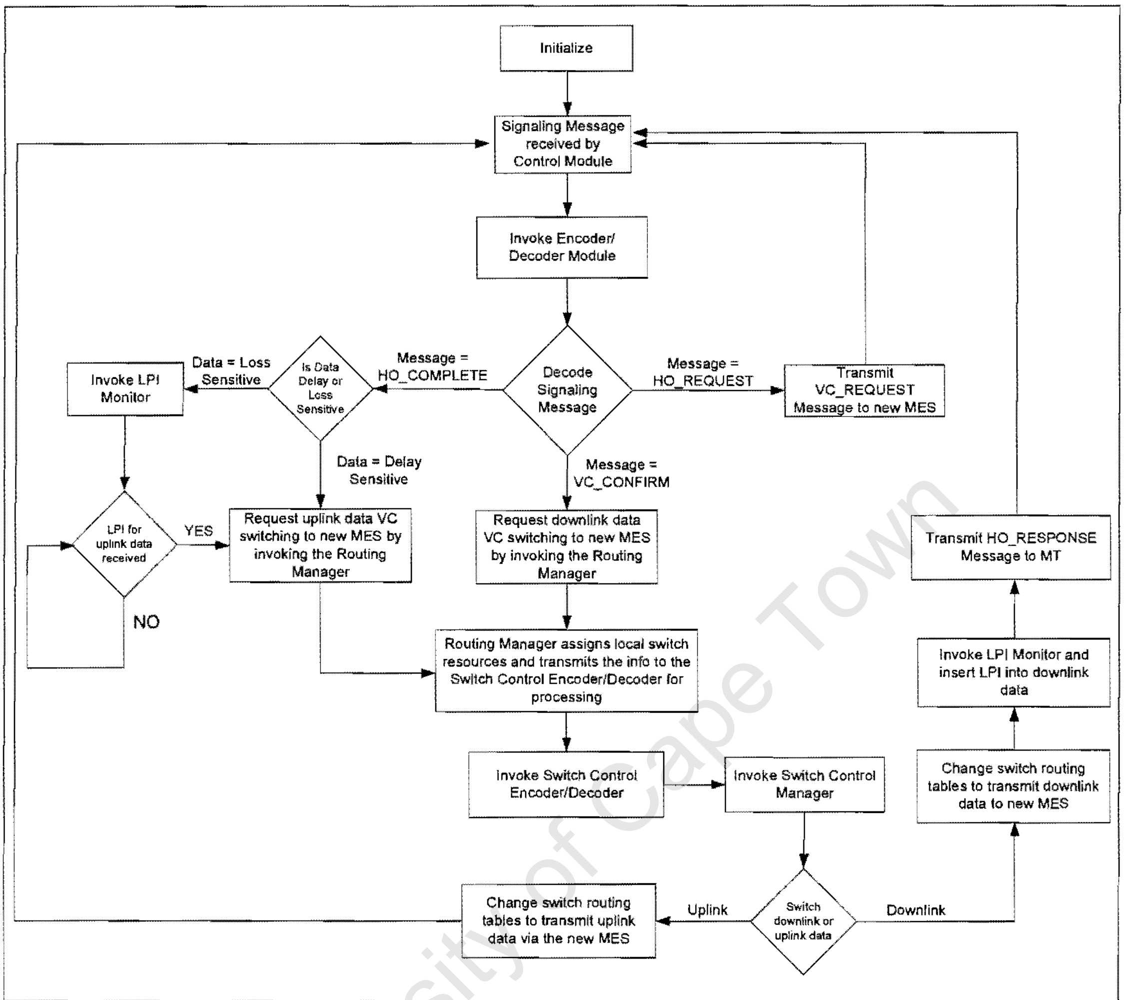


Fig. D4 MESSE handover flow diagram at the old MES

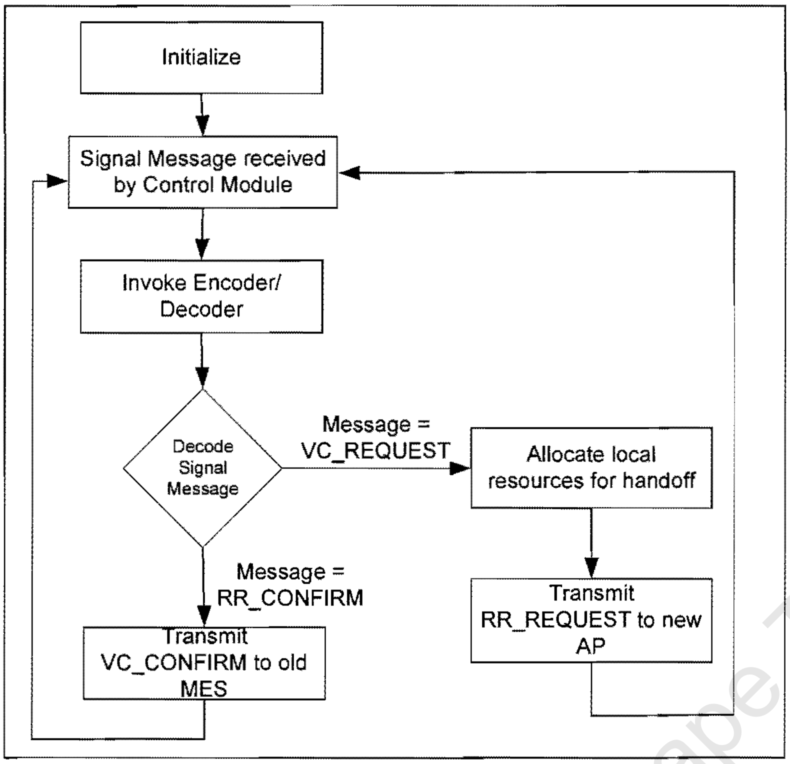


Fig. D5 MESSE handover flow diagram at the new MES

D.2 Handover Protocol Flow Diagrams for the PR Scheme

This section takes a closer look at the handover signaling processing in the various network elements for the PR scheme.

D.2.1 Mobile Terminal Signaling Entity

The MTSE is implemented as a software entity on the MT. The flow diagram for the handover protocol at the MT is the same as that of the PE scheme and was illustrated in Fig. D1.

D.2.2 Access Point Signaling Entity

The APSE is implemented as a software entity on the old and new APs. The flow diagram for the handover protocol at the old AP and new AP is also identical to that of the PE scheme and was illustrated in Fig. D2 and Fig. D3 respectively.

D.2.3 MES and COS Signaling Entities

The flow diagrams for the MES Signaling Entities at the old MES and the new MES for the PR Scheme is illustrated in Fig. D6 and Fig. D7 respectively.

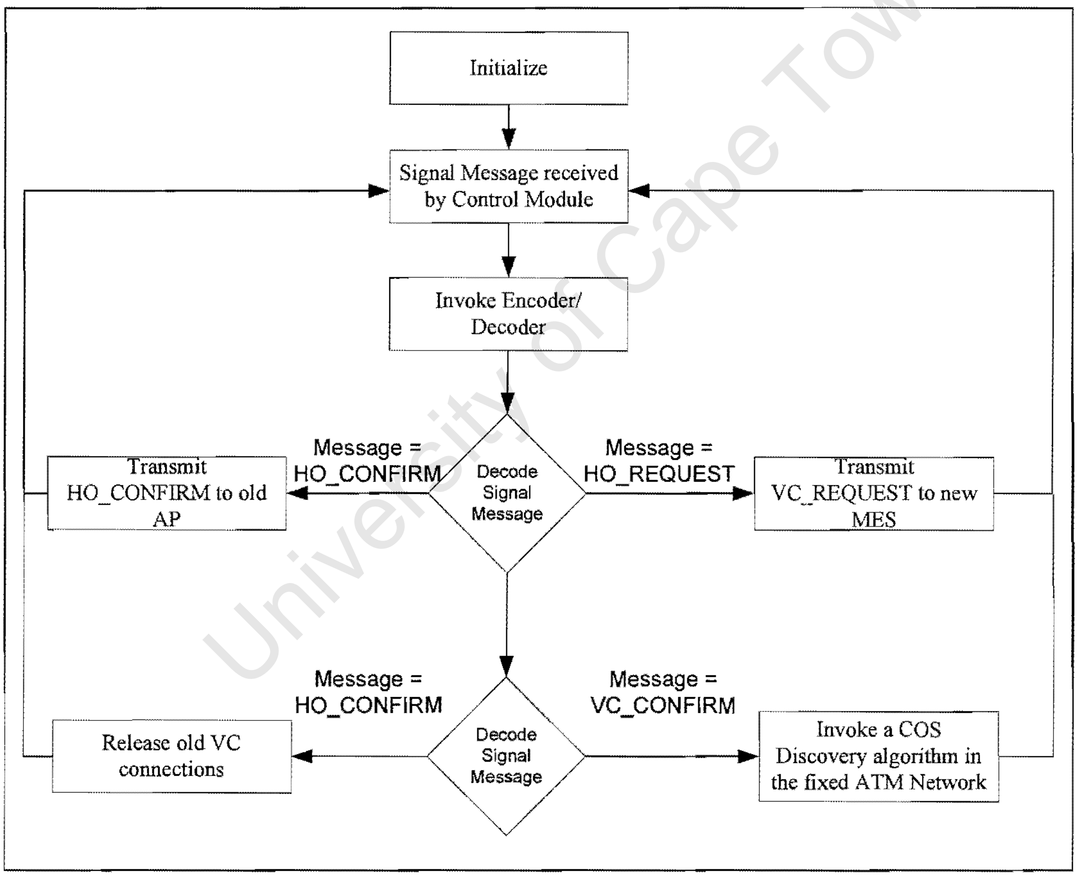


Fig. D6 MESSE handover flow diagram at the old MES

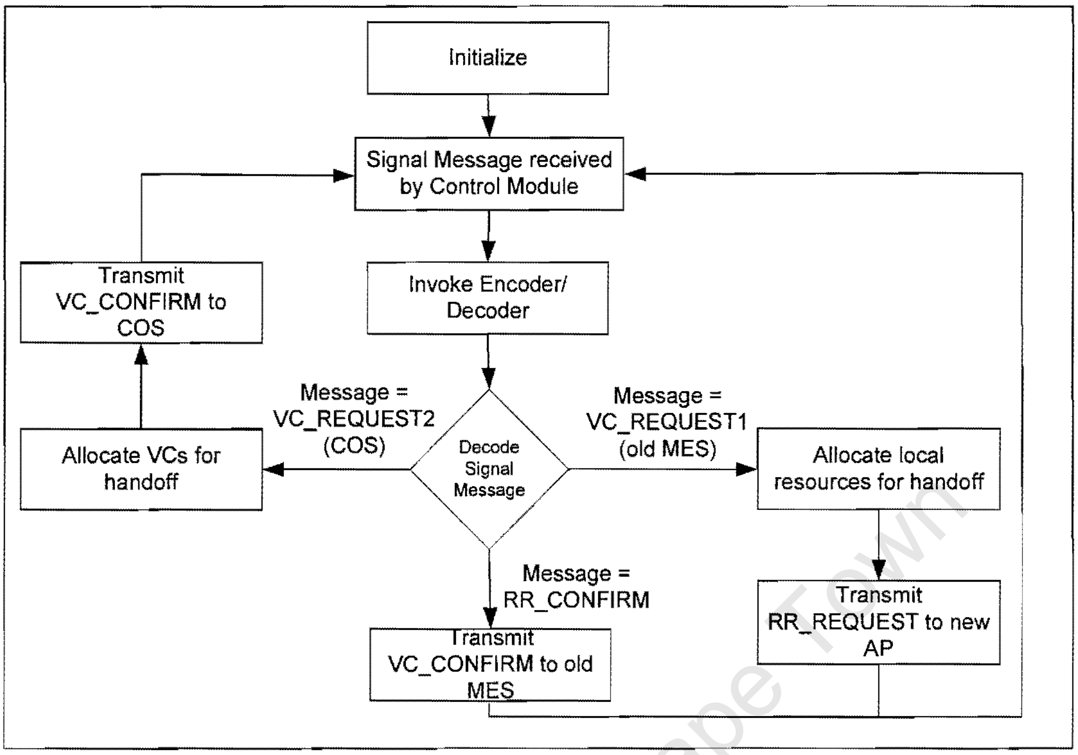
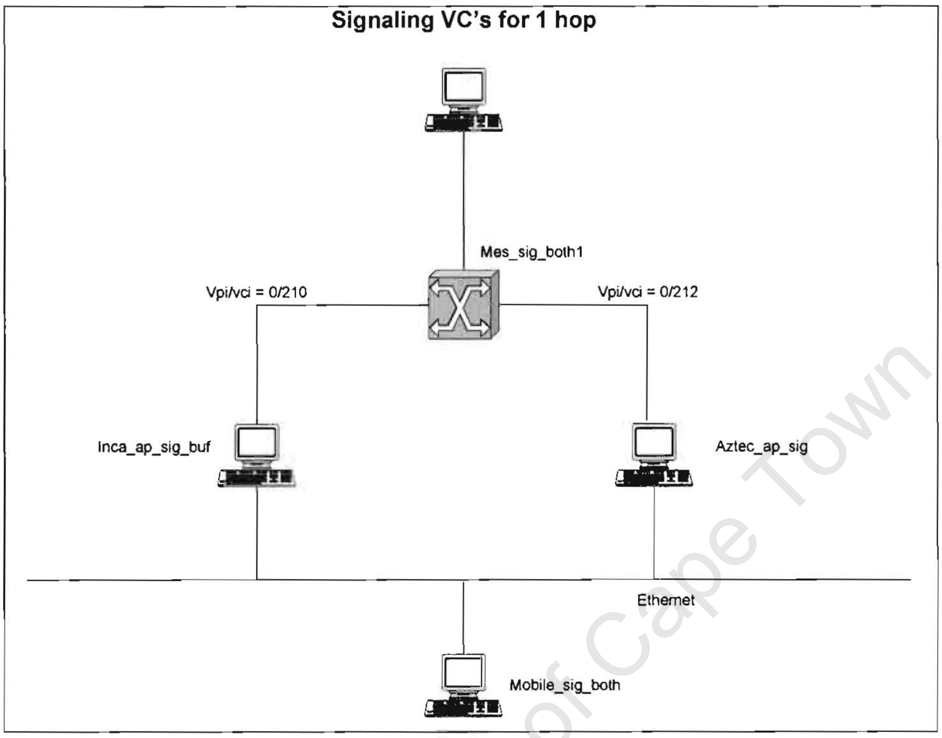


Fig. D7 MESSE handover flow diagram at the new MES

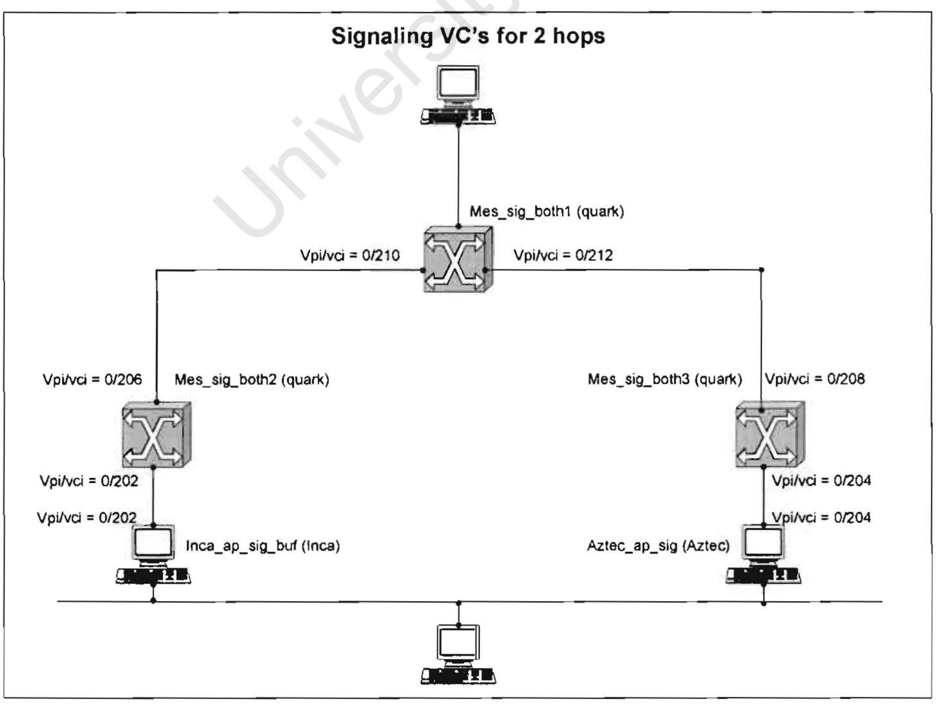
In the PR scheme, the functions of the MES are that of reserving its local resources for the handover and requesting radio resources from the new AP. In the PE scheme, the old MES was responsible for switching of the uplink and downlink data paths, whereas in the PR scheme, the COS is responsible for the switching of the data paths.

The handover flow diagram for the COS is identical to the handover flow diagram of the old MES in the PE scheme as illustrated in Fig.D4. The only difference is that the COS would not request radio resource allocation from the new AP.

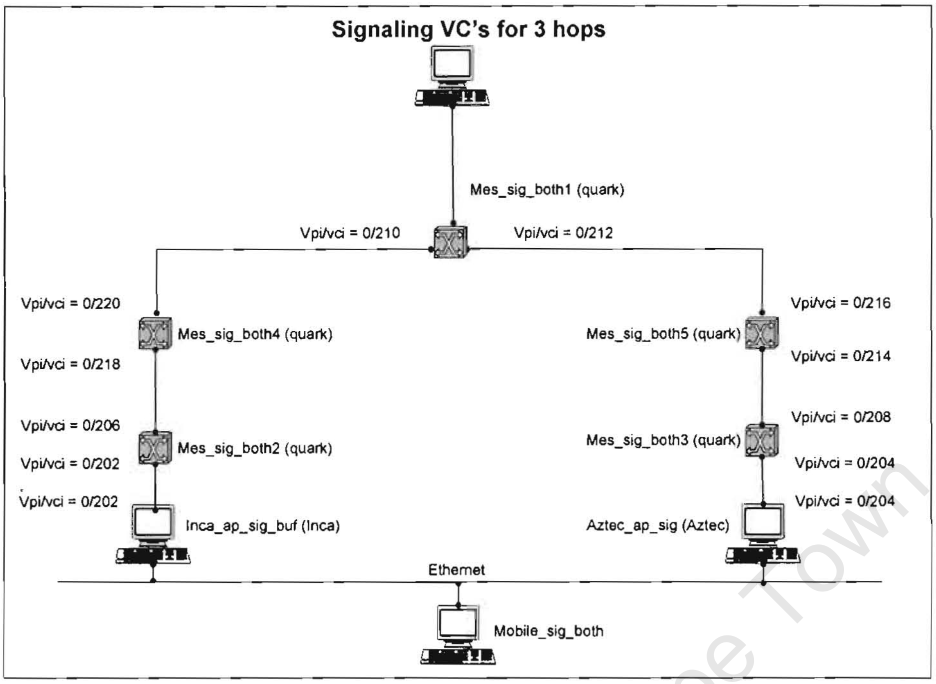
Appendix E



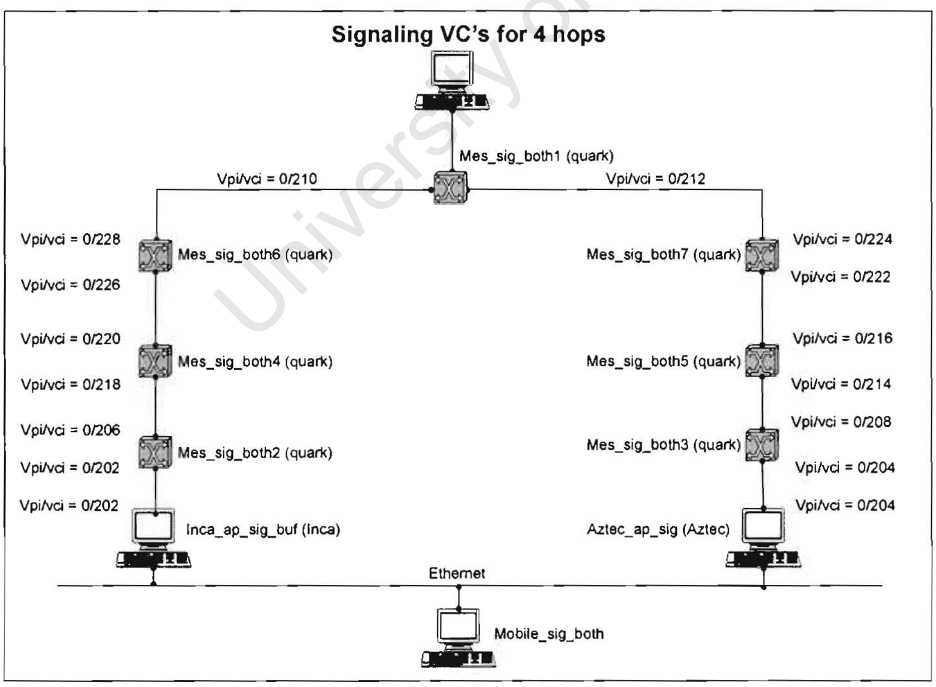
Signaling VC's for 1 hop
Mes_sig_both1
Vpi/vci = 0/210
Vpi/vci = 0/212
Inca_ap_sig_buf
Aztec_ap_sig
Ethernet
Mobile_sig_both



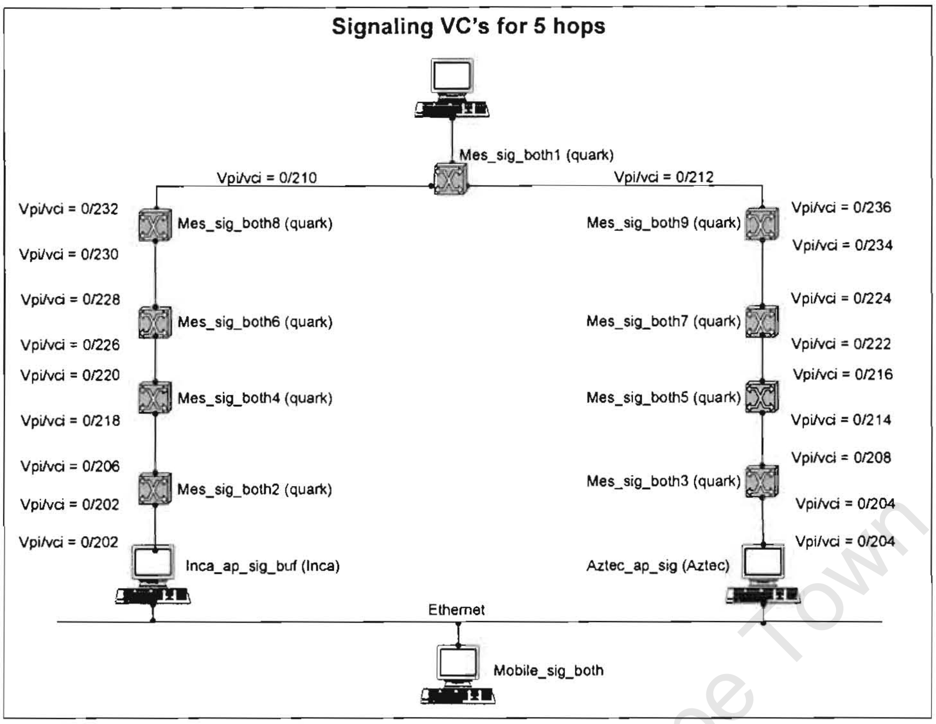
Signaling VC's for 2 hops
Mes_sig_both1 (quark)
Vpi/vci = 0/210
Vpi/vci = 0/212
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Mes_sig_both3 (quark)
Vpi/vci = 0/208
Vpi/vci = 0/202
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Vpi/vci = 0/204
Vpi/vci = 0/204
Aztec_ap_sig (Aztec)

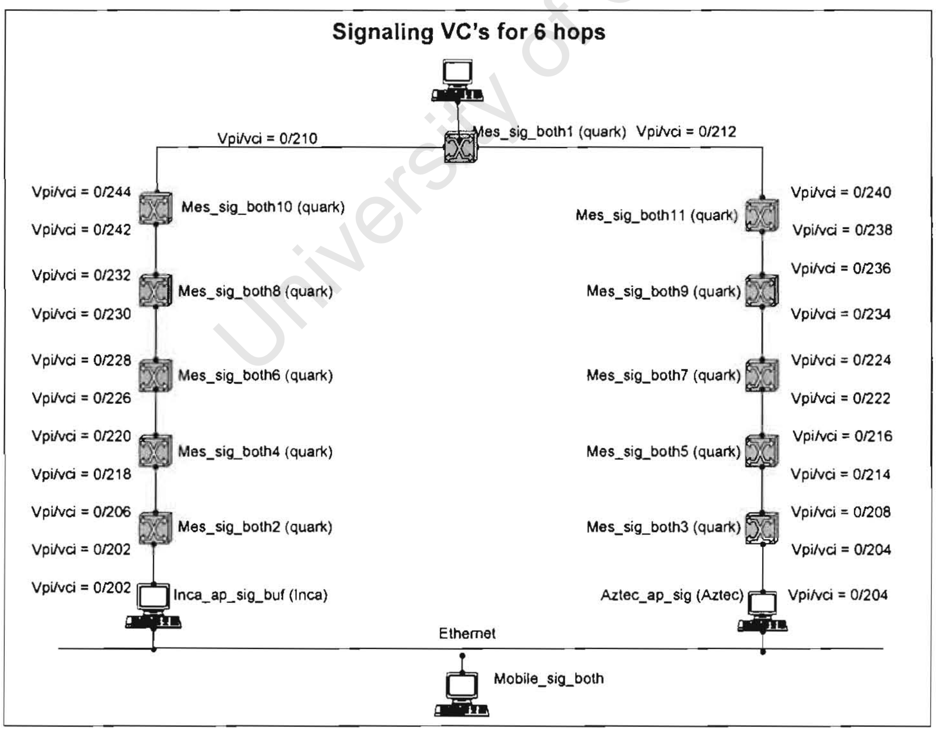


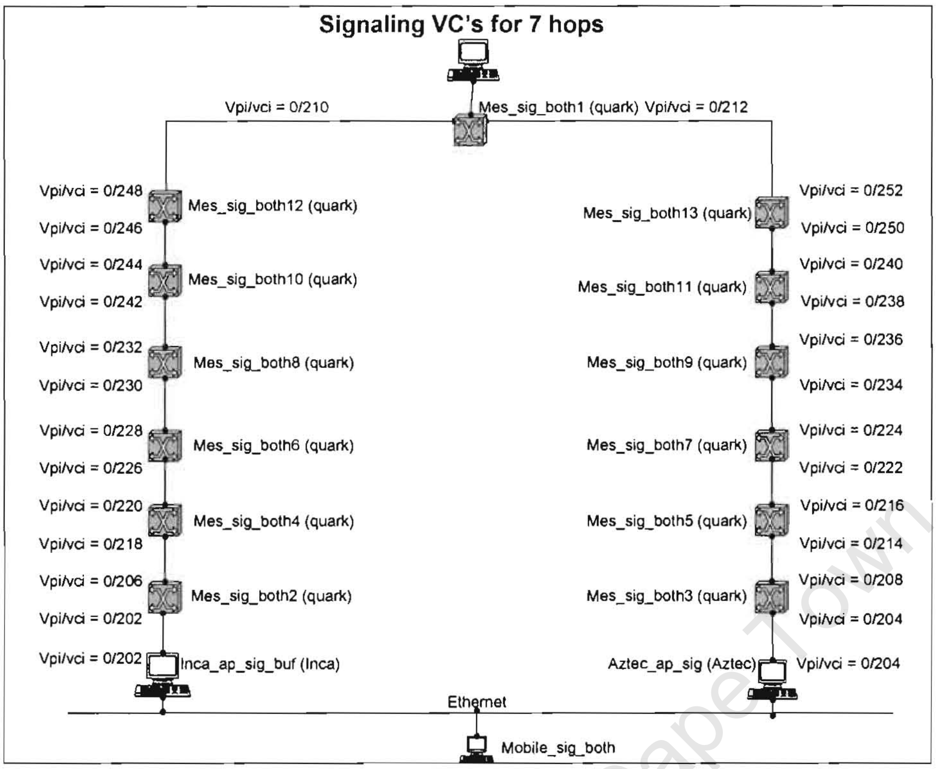
Signaling VC's for 3 hops
Mes_sig_both1 (quark)
Vpi/vci = 0/210
Vpi/vci = 0/212
Vpi/vci = 0/220
Mes_sig_both4 (quark)
Vpi/vci = 0/218
Vpi/vci = 0/216
Mes_sig_both5 (quark)
Vpi/vci = 0/214
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Vpi/vci = 0/202
Vpi/vci = 0/208
Mes_sig_both3 (quark)
Vpi/vci = 0/204
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Vpi/vci = 0/204
Aztec_ap_sig (Aztec)
Ethernet
Mobile_sig_both

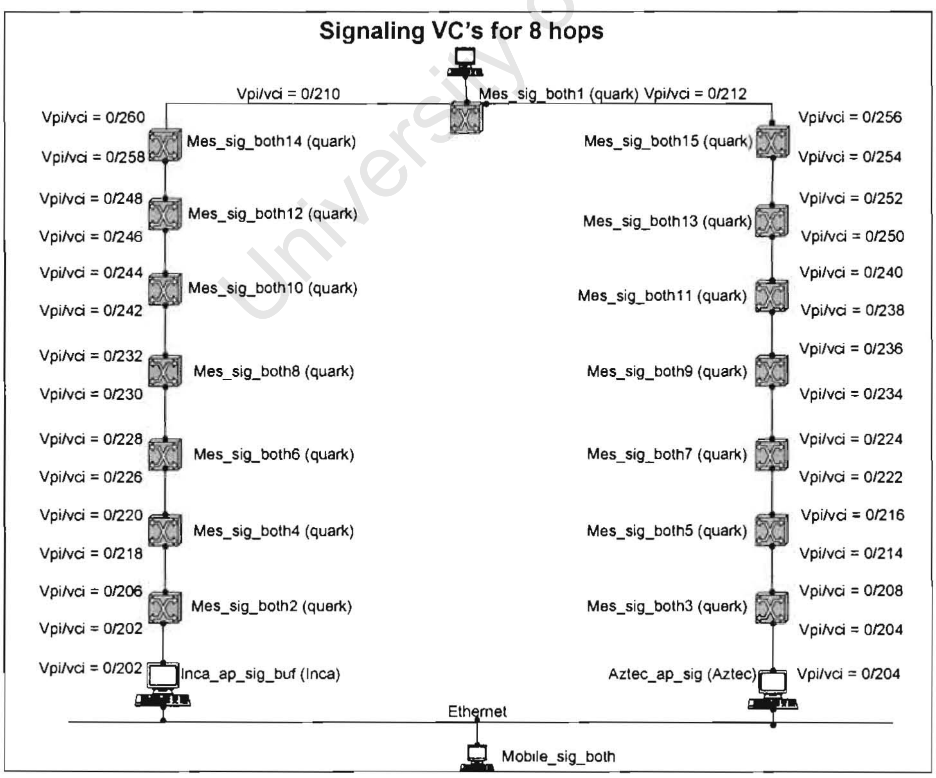


Signaling VC's for 4 hops
Mes_sig_both1 (quark)
Vpi/vci = 0/210
Vpi/vci = 0/212
Vpi/vci = 0/228
Mes_sig_both6 (quark)
Vpi/vci = 0/226
Vpi/vci = 0/224
Mes_sig_both7 (quark)
Vpi/vci = 0/222
Vpi/vci = 0/220
Mes_sig_both4 (quark)
Vpi/vci = 0/218
Vpi/vci = 0/216
Mes_sig_both5 (quark)
Vpi/vci = 0/214
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Vpi/vci = 0/202
Vpi/vci = 0/208
Mes_sig_both3 (quark)
Vpi/vci = 0/204
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Vpi/vci = 0/204
Aztec_ap_sig (Aztec)
Ethernet
Mobile_sig_both


Signaling VC's for 5 hops
Mes_sig_both1 (quark)
Vpi/vci = 0/210
Vpi/vci = 0/212
Vpi/vci = 0/232
Mes_sig_both8 (quark)
Vpi/vci = 0/230
Mes_sig_both9 (quark)
Vpi/vci = 0/236
Vpi/vci = 0/234
Vpi/vci = 0/228
Mes_sig_both6 (quark)
Vpi/vci = 0/226
Mes_sig_both7 (quark)
Vpi/vci = 0/224
Vpi/vci = 0/222
Vpi/vci = 0/220
Mes_sig_both4 (quark)
Vpi/vci = 0/218
Mes_sig_both5 (quark)
Vpi/vci = 0/216
Vpi/vci = 0/214
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Vpi/vci = 0/202
Mes_sig_both3 (quark)
Vpi/vci = 0/208
Vpi/vci = 0/204
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Aztec_ap_sig (Aztec)
Vpi/vci = 0/204
Ethernet
Mobile_sig_both


Signaling VC's for 6 hops
Vpi/vci = 0/210
Mes_sig_both1 (quark)
Vpi/vci = 0/212
Vpi/vci = 0/244
Mes_sig_both10 (quark)
Vpi/vci = 0/242
Mes_sig_both11 (quark)
Vpi/vci = 0/240
Vpi/vci = 0/238
Vpi/vci = 0/232
Mes_sig_both8 (quark)
Vpi/vci = 0/230
Mes_sig_both9 (quark)
Vpi/vci = 0/236
Vpi/vci = 0/234
Vpi/vci = 0/228
Mes_sig_both6 (quark)
Vpi/vci = 0/226
Mes_sig_both7 (quark)
Vpi/vci = 0/224
Vpi/vci = 0/222
Vpi/vci = 0/220
Mes_sig_both4 (quark)
Vpi/vci = 0/218
Mes_sig_both5 (quark)
Vpi/vci = 0/216
Vpi/vci = 0/214
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Vpi/vci = 0/202
Mes_sig_both3 (quark)
Vpi/vci = 0/208
Vpi/vci = 0/204
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Aztec_ap_sig (Aztec)
Vpi/vci = 0/204
Ethernet
Mobile_sig_both


Signaling VC's for 7 hops
Vpi/vci = 0/210
Mes_sig_both1 (quark) Vpi/vci = 0/212
Vpi/vci = 0/248
Mes_sig_both12 (quark)
Vpi/vci = 0/246
Vpi/vci = 0/252
Mes_sig_both13 (quark)
Vpi/vci = 0/250
Vpi/vci = 0/244
Mes_sig_both10 (quark)
Vpi/vci = 0/242
Vpi/vci = 0/240
Mes_sig_both11 (quark)
Vpi/vci = 0/238
Vpi/vci = 0/232
Mes_sig_both8 (quark)
Vpi/vci = 0/230
Vpi/vci = 0/236
Mes_sig_both9 (quark)
Vpi/vci = 0/234
Vpi/vci = 0/228
Mes_sig_both6 (quark)
Vpi/vci = 0/226
Vpi/vci = 0/224
Mes_sig_both7 (quark)
Vpi/vci = 0/222
Vpi/vci = 0/220
Mes_sig_both4 (quark)
Vpi/vci = 0/218
Vpi/vci = 0/216
Mes_sig_both5 (quark)
Vpi/vci = 0/214
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Vpi/vci = 0/202
Vpi/vci = 0/208
Mes_sig_both3 (quark)
Vpi/vci = 0/204
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Aztec_ap_sig (Aztec)
Vpi/vci = 0/204
Ethernet
Mobile_sig_both


Signaling VC's for 8 hops
Vpi/vci = 0/210
Mes_sig_both1 (quark) Vpi/vci = 0/212
Vpi/vci = 0/260
Mes_sig_both14 (quark)
Vpi/vci = 0/258
Vpi/vci = 0/256
Mes_sig_both15 (quark)
Vpi/vci = 0/254
Vpi/vci = 0/248
Mes_sig_both12 (quark)
Vpi/vci = 0/246
Vpi/vci = 0/252
Mes_sig_both13 (quark)
Vpi/vci = 0/250
Vpi/vci = 0/244
Mes_sig_both10 (quark)
Vpi/vci = 0/242
Vpi/vci = 0/240
Mes_sig_both11 (quark)
Vpi/vci = 0/238
Vpi/vci = 0/232
Mes_sig_both8 (quark)
Vpi/vci = 0/230
Vpi/vci = 0/236
Mes_sig_both9 (quark)
Vpi/vci = 0/234
Vpi/vci = 0/228
Mes_sig_both6 (quark)
Vpi/vci = 0/226
Vpi/vci = 0/224
Mes_sig_both7 (quark)
Vpi/vci = 0/222
Vpi/vci = 0/220
Mes_sig_both4 (quark)
Vpi/vci = 0/218
Vpi/vci = 0/216
Mes_sig_both5 (quark)
Vpi/vci = 0/214
Vpi/vci = 0/206
Mes_sig_both2 (quark)
Vpi/vci = 0/202
Vpi/vci = 0/208
Mes_sig_both3 (quark)
Vpi/vci = 0/204
Vpi/vci = 0/202
Inca_ap_sig_buf (Inca)
Aztec_ap_sig (Aztec)
Vpi/vci = 0/204
Ethernet
Mobile_sig_both

Appendix F

WATM Design considerations

F.1 Access Point Design Considerations

Currently GSM and mobile Internet Protocol (IP) are widely used technologies supporting mobile users. Both support handoff. WATM provides a high-speed backbone network that can support mobility in a common infrastructure network to different wireless access technologies. A generic WATM network supporting GSM, wireless local area networks (LANs), and wireless ATM access can be designed as described below. In such a WATM network, the ATM level connection is terminated at the AP, and the received traffic is forwarded on the wireless segment using a link-layer access method that is specific to the mobile technology being emulated on the WATM network. The AP thus acts as a gateway, converting traffic from a format specific to the wireless link to a suitable ATM Adaptation Layer (AAL) on the wired ATM link. For example, if an AP within the WATM network offers a wireless LAN, e.g. 802.11, an IP terminal may send IP packets to the AP by encapsulating the packets within (Ethernet-like) wireless LAN frames. The AP forwards the received packets onto an ATM connection. Additionally, the IP terminal may use the mobility support offered by WATM (instead of mobile IP) by mapping IP onto WATM: the terminal's migration between APs are transparent to IP, and WATM is used for handoffs (instead of mobile IP).

F.2 Signaling in the ATM-based B-ISDN

Relevant Protocols for ATM are defined within a generic protocol architecture, the B-ISDN protocol reference model [48], which contains orthogonal dimensions for different layers and various functions.

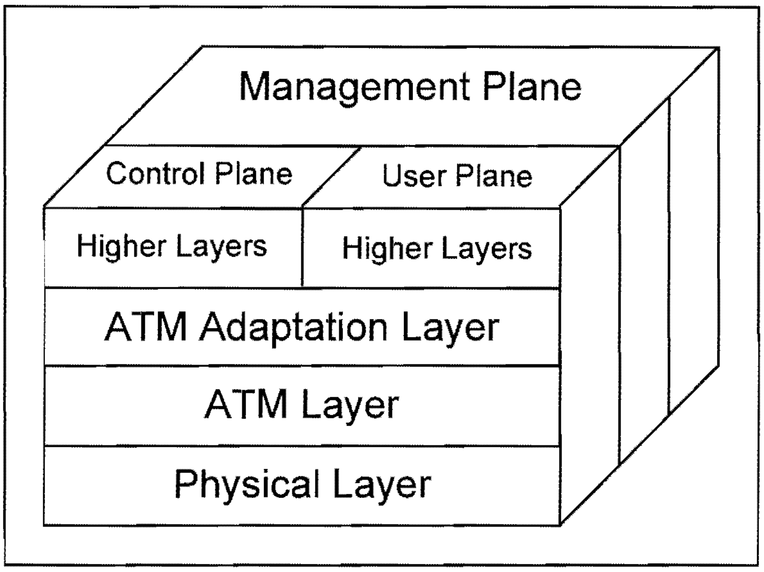


Fig. F.1 B-ISDN Reference Model

Horizontal layers encompass:

- Physical Layer - specifying media technology dependent issues
- ATM Layer – ATM specific functions, such as generation of ATM cell headers, the multiplexing de-multiplexing of cell.
- ATM adaptation Layer (AAL) – segmentation and re-assembling of higher layer protocol data units into ATM cells, and convergence functions.
- Higher Layers – for application specific functions.

The vertical structure defines the User Plane, the Control Plane and the Management Plane. User and control plane use identical physical and ATM layers respectively, but use different AAL layers. The user plane is responsible for transmission of user data, where the control plane is responsible for handling all relevant issues on signaling. Finally, the management plane encompasses relevant functions to interact and coordinate between user and control plane activities.

The WATM architecture design comprises of two distinct components, namely the ATM user-plane switching component and a handover control-plane signaling component. The user-plane switching component is standardized in ATM switching networks. The control-plane signaling component, which is the focus of the rest of this chapter, provides the necessary signaling functionality required to establish dynamic connections to facilitate a handoff in a WATM network.

F.3 Key Issues in ATM Signaling Scenario

A necessary prerequisite for transferring signaling messages is a reliable network service. Therefore, a Signaling ATM Adaptation Layer (SAAL) based on AAL 5 has been defined in ITU-T Recommendation Q.2100 [43]. Different types of AAL may be used as well, e.g., SPANS runs on top of either AAL3 or AAL 5 [44]. The ATM adaptation protocol has been divided into five types, namely AAL1, AAL2, AAL3, AAL4 and AAL5. However, AAL5 is the most flexible as it can support any type of service. Moreover, ATM signaling may be processed on top of the TCP/IP stack [45, 46], and may be used for transmitting signaling messages, to ensure the reliable transfer of these messages, as it is done by the Q.93B-based Netcomm signaling protocol [47]. AAL5 connections allow ATM networks to interface with the internets transport protocol, TCP/IP, by packing the IP packets into ATM cells. We have consequently chosen AAL5 as the transport protocol in the fixed ATM network.

Signaling protocols may also not require reliable signaling message transfer services at all, when the application itself is responsible for dealing with errors, such as corrupted, lost, or duplicated signaling messages [48]. In the Ethernet network, a choice exists between using TCP/IP or UDP/IP as the transport protocol for the signaling and data packets. TCP is connection-oriented, which means there must be a connection to setup before any data can be sent. TCP is also a reliable protocol, if the checksum detects there is a corrupted data, it will send back a request back to the sender for a retransmission. It also needs to send a confirmation to the sender when the data is received. If the data is lost during transmission, there will be another retransmission.

Unlike TCP, UDP is a connectionless protocol, which means that there is no connection between the sender and receiver, and the data will route individually at each intermediate node through the network. UDP therefore is also an unreliable protocol. UDP doesn't have any acknowledgment to notify the sender the data has been safely arrived. UDP therefore have a higher throughput compare to TCP, it sacrifice reliability for performance. Therefore, UDP was chosen as the transport protocol for the signaling and data traffic in the wireless (Ethernet) network. UDP was also chosen because a wireless channel is bandwidth limited. The 8 bytes header of UDP is more attractive and has less overhead than TCP's 20 byte header. The utilization of the bandwidth is therefore increased. Since, our transport need to be a high performance protocol, UDP was the best choice for data and signaling transfer applications.

F.4 Virtual Signaling Channels to support handover

In ATM-Based B-ISDN networks Out-of-Band signaling VCs are used [49]. They form the main basis for processing signaling tasks, e.g. transmitting and receiving signaling control messages. In-band signaling can also be utilized, however, utilizing in-band signaling to transport signaling messages with user data would require each ATM switch, AP and MT to interrogate each traversing ATM cell to determine whether it contains signaling information. This method would be time consuming. Out-of-band signaling VCs are faster at processing of signaling messages as a separate ATM VC for signaling is utilized.

In general, three different types of Signaling Virtual Channels (SVC) are distinguished [48]:

- Meta Virtual Signaling Channel (MVSC)
- Broadcast Virtual Signaling Channel (BVSC)
- Point-to-Point Virtual Signaling Channel (PVSC)

MVSCs are always bi-directional and are used to establish BVSCs or PVSCs as necessary. These in turn, are used to signal all types of ATM connection signaling messages between different ATM end –systems and ATM switches. BVSCs are uni-directional, while PVCSs are bi-directional. At least one VSC of either type is mandatory and therefore permanent for each single User to Network Interface (UNI). However, one static VSC, regardless of the type, has to be assigned at a defined (pair of) VPIs and VCIs initially for each ATM end-system. For PVSCs the VCI is always set to 5 and the VPI is always non-zero, while a MVSC is defined by the VCI set to 1 and the VCI equal always non-zero, if signaling is done with remote networks, else if the VCI equals zero, user signaling is processed at the local switch [50].

In the handover signaling design, PVSCs are utilizes to carry signaling messages. However, the VCI is not always set to 5. Any value can be used for the VPI/VCI of the signaling channel. If the VPI/VCI is not set to the values specified by the standards [50], it simply means that the signaling messages will be routed to propriety signaling entities that are specifically waiting for signaling messages on those VCs.

The network provides the PVSCs for signaling purposes and these circuits are ready for use by the signaling entities, e.g. MT, AP, MES etc.

Appendix G

G.1The WUGS Architecture

The WUGS is a high speed, multicast virtual circuit experimental ATM switch developed at Washington University [35, 36]. The open architecture of the WUGS enables experimental modification at all levels. The switch's external cell format follows the ATM standard and therefore the switch can be integrated with ease into local and wide area networks with little or no modification needed. One feature in particular, is that the switch has no internal processing capability, and therefore requires an external controller in the form of a PC.

The standard WUGS environment consists of the following components:

- The ATM Switch
- The ATM Port Interconnect Chip Network Interface Card (APIC NIC)
- Line card adapters.

The APIC NIC allows end stations to communicate with the switch over an optical link while the line card adapters provide the optical to electrical conversions.

We focus in this section on the WUGS environment that will facilitate the implementation of the MES.

G.2.1 The WUGS Switch Overview

The WUGS is an 8-port ATM Switch that allows connectivity from an optical medium to the switch port processors via two separate line cards. A dual OC3- line card allows per port throughput of 310 Mbits/s (2 x 155 Mbits/s) and a line card operating at a much higher speed of 1.2 Gbits/s is also available. The latter is called the APIC NIC.

Consider Fig. G.2.1 which illustrates the internal configuration of the WUGS. The switch fabric forms the core of the switch and connects each Input Port Processor (IPP) to all possible Output Port Processors (OPP). The IPP receive cells from the incoming links, buffer them while awaiting transmission through the switching fabric and perform the virtual path/circuit translation required to route cells to their proper output (or outputs). The OPP resequence cells received from the switching fabric and queue them while they await transmission on the outgoing link. Each OPP is also connected to its corresponding IPP, providing the ability to *recycle* cells belonging to multicast connections.

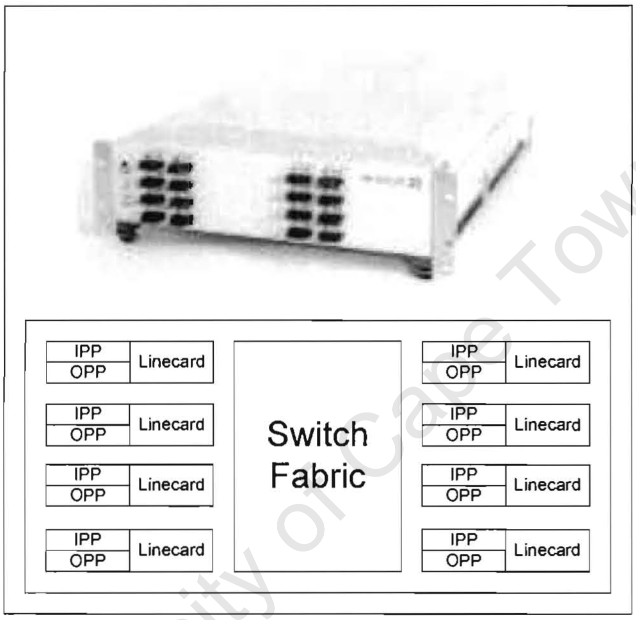


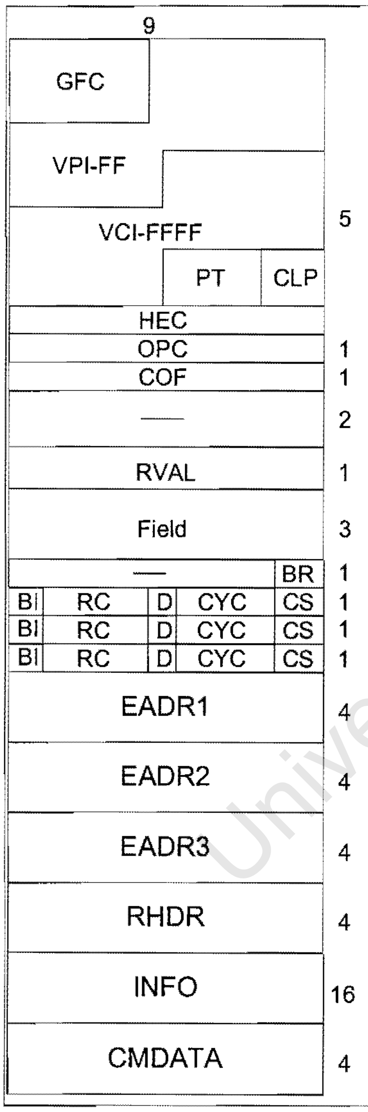
Fig. G.1 Internal organization of the WUGS

G.2.2 Gigabit Network Switch Controller (GBNSC)

The purpose of this software is to control the WUGS and hide hardware details as much as possible [37]. GBNSC monitors the state of the switch and provides access to all hardware details for client applications. Since the switch has no processing capability, a standard PC running GBNSC controls the switch. Access to the switch is through ATM cells transmitted by the controller. These special cells, called control cells, have special formats defined and are sent on VPI 0 VCI 32. The switch's internal routing tables and maintenance registers are modified and monitored by the control cells.

GBNSC operations are initiated by two types of events, the receipt of a request message from another process (Jammer) and internal events (e.g. time-outs).

Sixteen operations are supported by the WUGS through control cells. The control cell formats are shown in Fig. G.2. The opcodes which may be used in the control cells are given in Table G.1. A full definition of all the fields can be found in [38].



Field	Description
ATM Header	Standard ATM Header Fields
OPC	Opcode
COF	Control Offset-count for which switch chip is to perform control operation.
RVAL	Return Value – code describing success or failure operation
FIELD	Target table entry
BR	Bypass Resequencer – for immediate release of cells at output port
BI	Busy/Idle – distinguish between idle and control or data cells.
RC	Routing Control – specific path, Port1, Port2, both or range copies.
D	Data Cell – distinguish between control and data cells.
CYC	Recycling
CS	Continuous Stream
EARDi	External Routing Addresses
RHDR	Return Header – for routing control cell back to CP.
CMDATA	Connection Management Data – used by CP to identify returning control cells.
LT	Local Time – local switch cell clock. Can be used by CP to compute traffic statistics.

Fig. G.2 Control Cell Fields

Opcode	Command	Description
0	NOP	No operation
F0 (Hex)	RST	Hard Reset of all chips
2	CLRERR	Clear all error flags in all chips.
3	RDVPXT	Read virtual path entry from VXT
4	RDVCXT	Read virtual circuit table entry from VXT
5	RDVPXTCC	Read cell counter (CC) from virtual path table.
6	RDVCXTCC	Read cell counter (CC) from virtual circuit table.
7	WRVPXT	Write virtual path table entry into VXT.
8	WRVCXT	Write virtual circuit table entry into VXT.
9	WRVPXTTR	Write virtual path table entry into VXT and start transitional time stamping.
10	WRVCXTTR	Write virtual circuit table entry into VXT and start transitional time stamping.
11	WRVPXTCC	Write cell counter (CC) to virtual path table (for testing only)
12	WRVCXTCC	Write cell counter (CC) to virtual circuit table (for testing only)
13	ERRORS	Return a cell only if error conditions exist.
14	RDMR	Read maintenance register field.
15	WRMR	Write maintenance register field.

Table G.1 Control Cell Opcodes

G.2.3 Jammer

Jammer is a script based client utility used to access all the bits in the WUGS tables and registers [39]. It connects to GBNSC through a TCP/IP socket and issues pre-defined commands to ping the switch, read or write routing tables, read maintenance registers, or reset/clear the switch's tables and registers. Users can create Jammer scripts to automate routing table programming. It is this feature of Jammer and GBNSC that was used to access the routing tables of the switch directly and modify it, in order to facilitate a fast handoff dynamically.

To understand how Jammer works, it is important to understand the communication structure shown in Fig. G.3.

Jammer provides the user full and direct access to the switch controller via the Node Communications Control Protocol (NCCP) [40]. Direct access to all registers and tables in the switch is provided to the user in this manner.

The Jammer syntax is a loose adaptation of the C and shell language syntaxes. To provide a certain degree of simplicity, some of the C syntax was modified and certain additions were made. Therefore, programmer's familiar with C programming has little difficulty adapting to writing Jammer scripts. For experienced programmers, the most difficult part is to understand how Jammer together with GBNSC controls the switch at the lower levels.

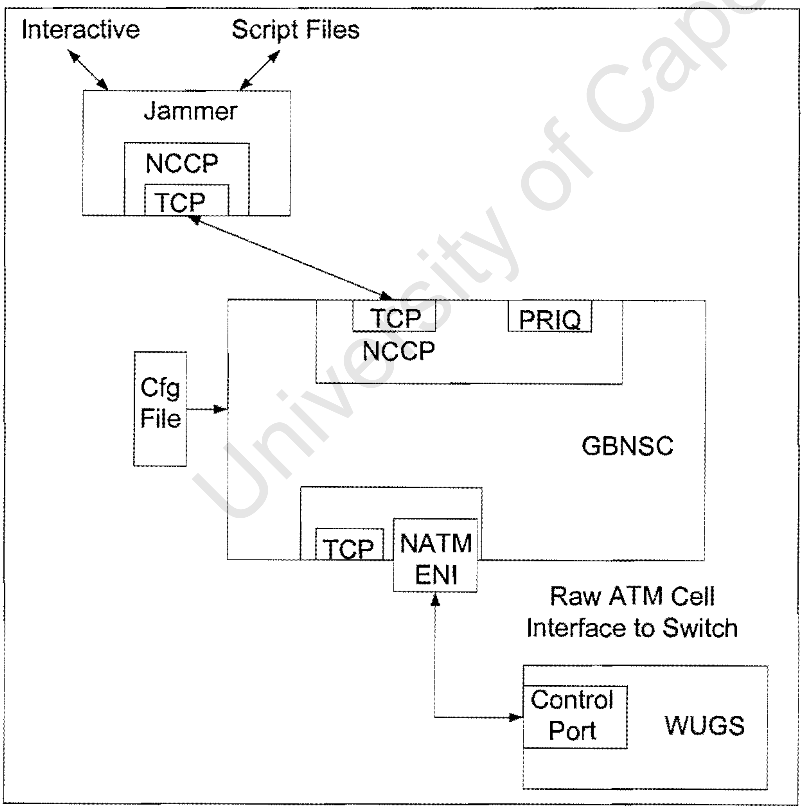


Fig. G.3 WUGS Software Overview

The following commands are available for manipulating routing tables and maintenance registers within the WUGS.

Command	Description
Reset sm	Reset the switch
Write mr Write vcxt Write vpxt	Commands for writing entries into maintenance registers, virtual circuit tables and virtual path tables.
Read mr Read vcxt Read vpxt	Commands for reading entries from the maintenance registers, virtual circuit tables and virtual path tables.
Clear errors	Clear all the error bits in all the maintenance registers.
Clear vpxt Clear vcxt	Clear a particular vcxt or vpxt entry.
Clear vxt	Clear ALL vcxt and vpxt entries at a port
Build Merge	A group of commands that make it easier to build multipoint recycling connection trees.

Table G.2 Switch Manipulating Commands

By using these commands in the Jammer utility, one can fully modify the routing tables and switch registers in order to facilitate a handoff. Fig. G.4 gives a graphical illustration of a virtual circuit read operation.

Each Input Port Processor (IPP) to the WUGS has a Virtual Translation Table (VXT) to direct the IPP what to do with received cells. Each VXT is divided into two separate areas, the Virtual Path Translation Table (VPXT) and the Virtual Circuit Translation Table (VCXT).

When an IPP receives a cell, it reads the Virtual Path Identifier (VPI) and Virtual Circuit Identifier (VCI) from the cell header. The IPP matches these values with the indexes of records in the VPXT and VCXT. Fields within the matching VPXT and VCXT instruct the IPP what to do with the cell. The IPP forwards the cell with control information to an Output Port Processor (OPP) based on these instructions. The user or network management system controls the flow of cells through the WUGS by manipulating the VXT in each IPP. By making use of the same process, the user can do similar write operations.

The fields that are manipulated by the network management to change the VXT entries are as follows.

By changing the VPXT and VCXT entries of the cells as they flow through the switch, we are able to switch connections at the ATM level as the handoff procedure occurs. This operation ensures a fast and efficient handoff as required by the initial set of requirements.

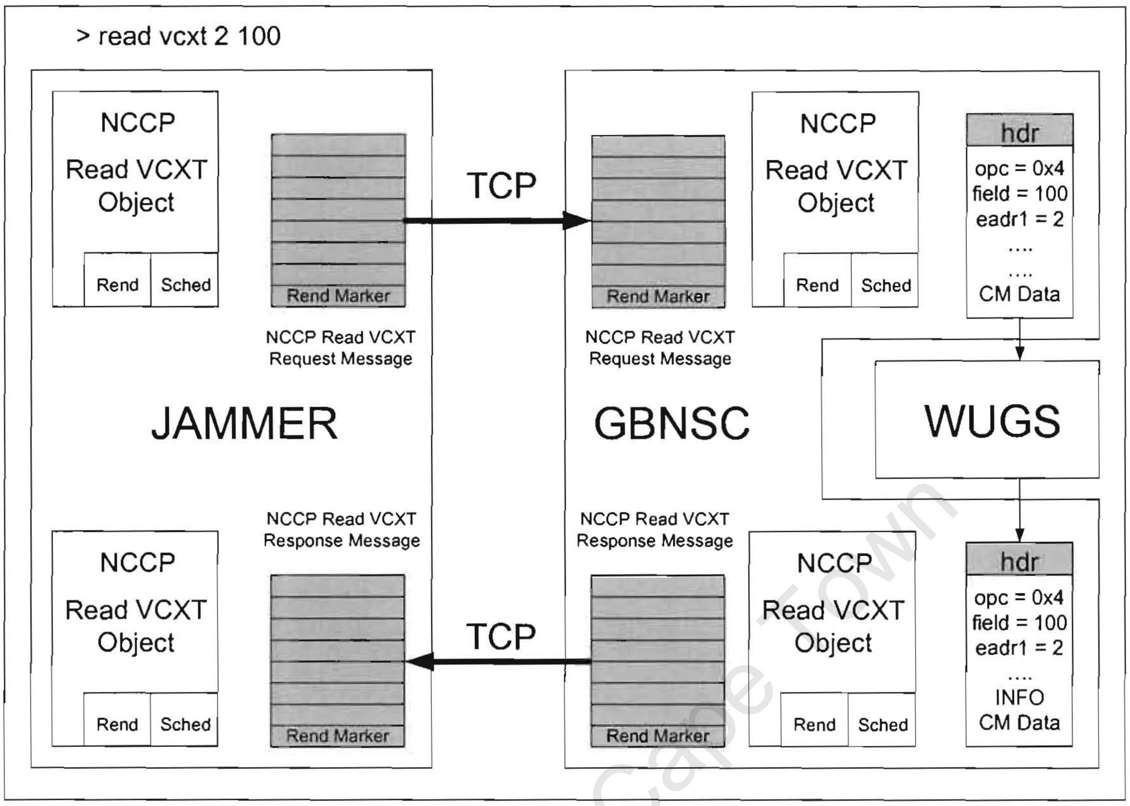


Fig. G.4 Reading a VCXT Entry

Appendix H

Step by Step Implementation Setup

A simplified subset of the logical WATM test bed shown in Fig 5.1 is illustrated in Fig. H.1. The first step in the implementation is to establish a signaling connection between a two ATM endpoints through the ATM switch. This figure illustrates the communication between the FH and one of the APs or the communication between two APs via the switch.

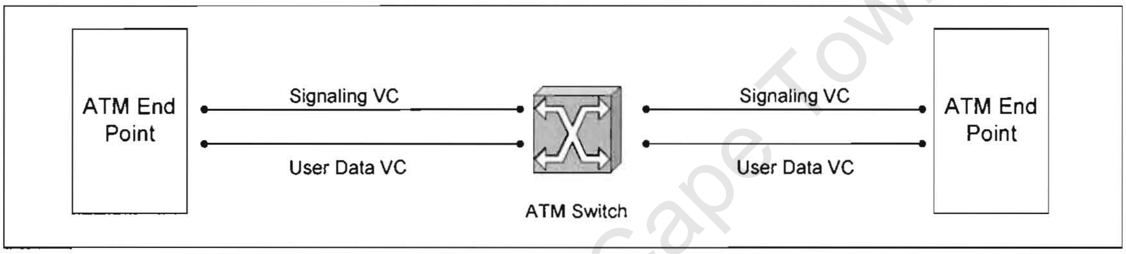


Fig. H.1 Network Architecture for a simple ATM end-to-end application

The protocol stack for this implementation is shown in Fig. H.2.

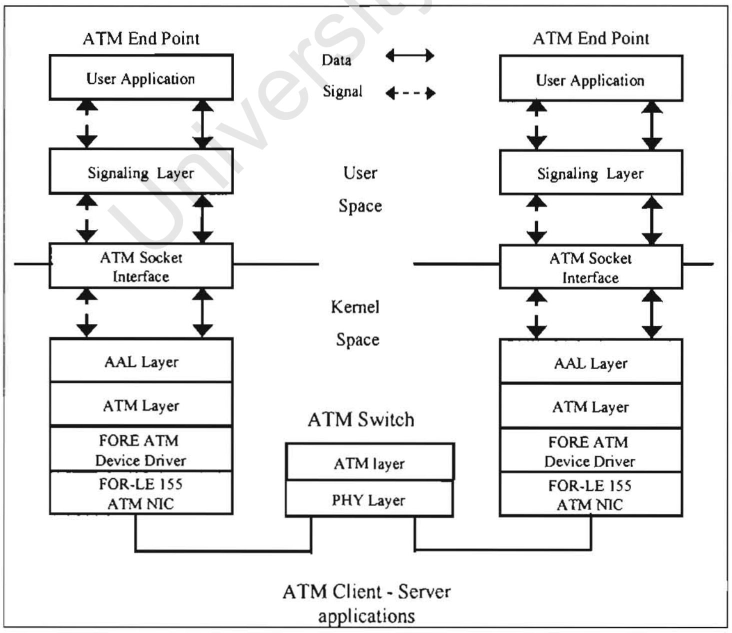


Fig. H.2 Protocol Stack for a simple ATM end-to-end application

The two ATM End Points are connected via multimode fibre to the WUGS switch ports with the ForerunnerLE ATM NIC (ENI155P).

The WATM test bed is implemented on the Linux operating system (Kernel 2.4.18). The Linux operating system is not ATM aware and therefore software have to be added before the operating system becomes ATM aware.

The ATM on Linux provides support for ATM networking under a traditional Linux-based operating system. The networking subsystem of the Linux kernel is composed of three layers: the socket layer, the protocol layer, and the network interface layer. Transmitted data travels from the application through the socket and protocol layers to the network interface layer. Received data arrives at the network interface and are passed up towards the socket layer. All data in the networking subsystem are stored in a data structure called an "mbuf". There are two basic types of mbufs: small mbufs and large mbufs. Small mbufs contain 108 bytes of data and are used for small data or packet headers. Large mbufs typically contain either 2K or 4K of data. Mbuf structures can be linked together to form an mbuf chain.

The socket layer has two main roles. First, it transfers the data between a user's address space and kernel layer mbufs. Second, it queues the data between the user and the kernel. If a process attempts to transmit too much data and its socket buffer becomes full, the socket layer will put the process to sleep until room is available.

All networking protocol processing is done at the protocol layer. ATM, TCP, UDP, and IP are all implemented in the Linux networking subsystem's protocol layer. When transmitting, the protocol layer receives data from the socket layer, adds the necessary headers, and passes the packet to the network interface layer for transmission. When receiving, the protocol layer dequeues packets from its input queue, and determines the destination of each packet. If the packet is to be forwarded to another host, then the protocol passes it back to the network interface layer. If the packet is bound for a local process, then the protocol layer enqueues the packet on the receiving process' receive buffer and notifies the socket layer that new data are available.

The network interface layer transfers packets between the networking hardware and the protocol layer. When transmitting, the network interface layer receives packets through its interface queue and transmits them on the network. When receiving, the network interface layer determines which protocol to pass the inbound packet to, enqueues the packet for the protocol, and then schedules a software interrupt to service the protocol.

ATM networking is integrated into the Linux kernel through a device-independent ATM networking layer and a device-specific driver for the ForerunnerLE ATM card (ENI155P). The device-independent layer provides support for using IP over ATM through PVCs and also provides support for "native" mode ATM sockets to send and receive raw ATM cells or AAL5 frames. An ATM pseudo header structure is used to route the ATM packets through the Linux networking subsystem. This four-byte header consists of the virtual circuit number (VPI and VCI) and a set of flag bits. The first flag bit indicates whether AAL0 or AAL5 is being used. This pseudo header is needed because the normal ATM header is removed from each cell in hardware by the network interface layer. The pseudo header only exists in the protocol layer and is removed before it is passed up to the socket layer or down to the network interface layer.

The device-dependent layer of ATM on Linux supports the ForerunnerLE 155Mbits/s ATM Network Card. To transmit data, the protocol layer enqueues an mbuf chain on the network device's input queue and calls the device's start routine. The start routine immediately removes the outbound packet from the network interface queue and inspects the packet's ATM pseudo header to determine on which transmit channel to enqueue the packet. Then the driver inserts a Transmit Buffer Descriptor (TBD) at the front of the packet and a trailer to the end of the data area so that it is the proper length. The TBD is read by the hardware to determine size and destination of the packet and then discarded. When the card receives a complete AAL5 frame or an AAL0 cell into its on-board memory it puts the virtual circuit on a hardware-managed service list and generates a receive interrupt. The driver's interrupt handler responds by taking the virtual circuit off the hardware service list and placing it on a software managed service list. The software list is needed in case there is a shortage of memory resources. The driver allocates mbuf chains for each frame and then programs the card to transfer the data from on-board memory to host memory. The mbuf chain receiving the data is placed on the receive queue and the driver removes the circuit from the software service list. The packet is pulled off the receive queue and passes it up to the protocol layer.

In order to describe how signaling entities implement sockets, we show how two machines A and B use the socket library to establish a bidirectional connection between them.

1. Machine A must first create a socket s and prepare the socket for receiving connection requests from machine B. This is done by invoking `listen(s)`.

2. A can now invoke `accept(s)` to block waiting for a connection request.

3. B creates a socket b and issues a connection request to machine A by invoking `connect`.

4. The connection request from B unblocks the `accept()` in A and creates a new socket a in machine A.

A bidirectional connection now exists between the sockets a and b allowing the machines A and B to communicate using the socket `send` and `recv` routines. Internally, the sockets a and b both consist of a *transmitter* and a *receiver*. The transmitter of socket a will transmit to the receiver of socket b and vice versa. Each receiver has a receiver buffer and an interrupt flag. Each transmitter has an interrupt flag which is used for flow control. The receiver will stimulate the transmitter interrupt flag when room becomes available in the receiver buffer. This allows a transmitter to pause transmission and go to sleep on the transmitter interrupt flag when the target receiver buffer runs full. The transmitter detects a full target buffer by inspecting the read pointer in the receiver buffer.

The second step in the implementation is to extend the connection from one ATM End Point to a MT. This extension will be made to the MT via an Ethernet network in order to simulate the wireless network. This is illustrated in Fig. H.3.

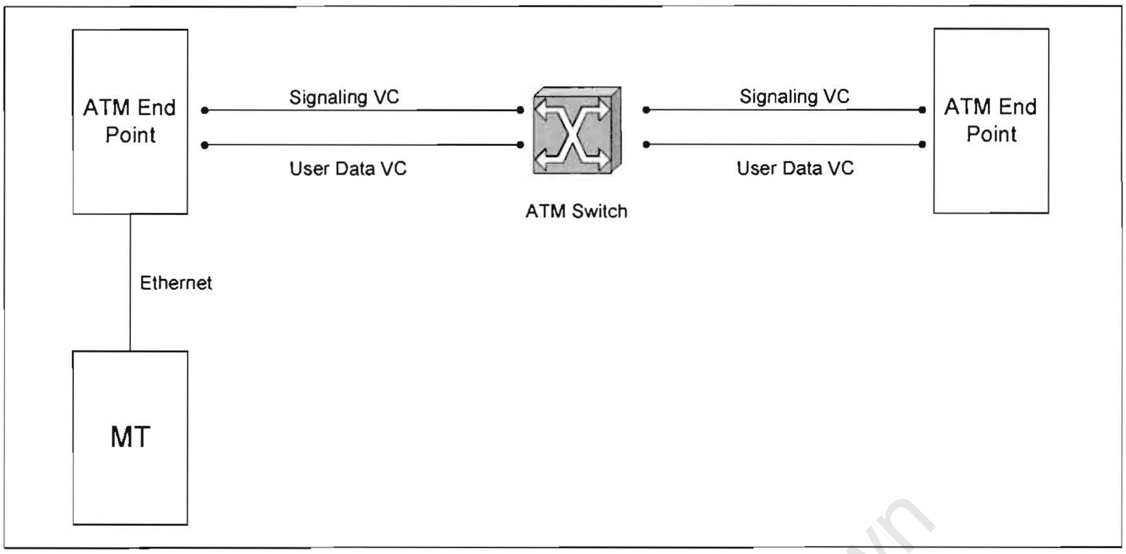


Fig. H.3 Network Architecture for connecting Ethernet to an ATM network

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 Mobile Terminal connected to the AP's via Ethernet. This would allow us to focus on the consequences of connection rerouting: loss, duplication, and reordering of packets.

The protocol stack for the implementation in Fig. H.3 is illustrated in Fig. H.4.

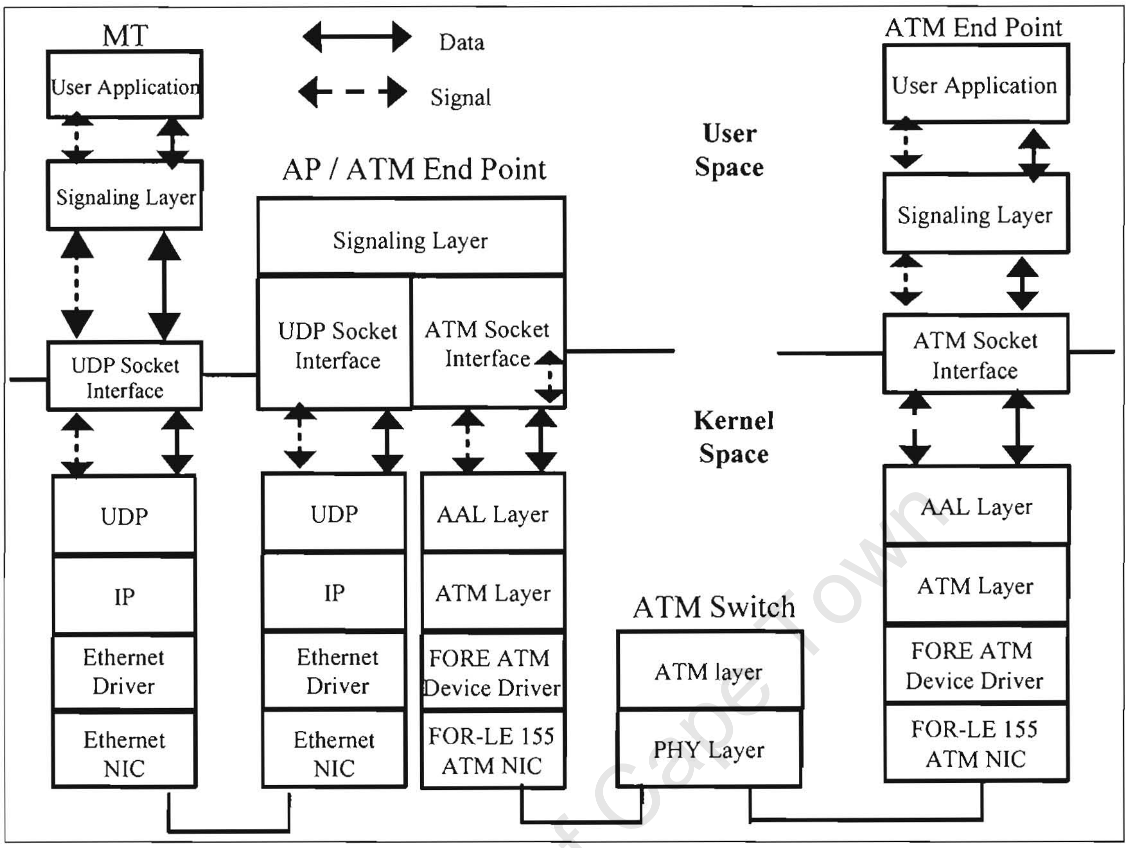


Fig. H.4 Protocol Stack for connecting Ethernet and ATM

The protocol stack for the Ethernet to ATM connection is identical to the previous one, except that we have the added Ethernet section. The ATM End Point 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 OC-3 dual card does not allow TX/RV on same VCI. Thus, the Tx of Lepton is connected to port 0A and its Rx is connected to port 1A. Conversely, the Rx of Quark is connected to port 0A and its Tx is connected to port 1A. The Jammer script file for creating uni-directional and bi-directional connections in the WUGS is given in Appendix C.