This is to certify that the

dissertation entitled

A Framework for Handling User Mobility in Wireless ATM Networks

presented by

Hui-Tang Lin

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Electrical Engineering

Major professor

Date 05-14-98

A Framework for Handling User Mobility in Wireless ATM Networks

By

Hui-Tang Lin

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1998

N

A

user

netw

In th

dard.

three

mana

T

the A

Abstract

A Framework for Handling User Mobility in Wireless ATM Networks

By

Hui-Tang Lin

Among issues related to the emerging wireless ATM (WATM) network, handling user mobility is considered especially a challenging. Handling user mobility on WATM networks involves two major issues: location management and handoff management. In this research, the focus is to provide a framework that extends the ATM standard, for handling user mobility on the future WATM network. This framework has three components: a virtual path WATM architecture (VPWA), a VPWA location management scheme, and a VPWA handoff management scheme.

The VPWA architecture is designed to allow constructing the WATM network on the ATM backbone network without major modifications on the deployed systems.

Its
net
era
bui
is i
able
by l
and
locat
overl
tion
fast a
unde
cells

T
First
metri
to ev
testb
enhar
funct
exper

Its infrastructure consists of an ATM backbone network and a wireless peripheral network. Most mobility desired functions are implemented on the wireless peripheral network. This avoids major modifications on the deployed ATM network when building the WATM network in the future. The VPWA location management scheme is introduced to provide the mobile location tracking function, which is not available on the ATM standard. This scheme complements the ATM routing mechanism by handling all mobile related calls. In conjunction with the VPWA architecture and several techniques, the VPWA location management scheme improves the IS-41 location management mechanism by reducing the excessive signaling and database overhead. The VPWA handoff management scheme applies an incremental connection re-establishment approach superimposed onto the VPWA architecture to provide fast and efficient connection reroutings. A special mechanism is introduced to reroute undelivered and misrouted cells to their destination. By doing so, the number of lost cells due to connection handoffs is reduced.

The proposed user mobility framework is evaluated by using two approaches. First, an analysis is conducted to investigate its theoretical performance. Performance metrics, such as the connection disruption time and buffer requirement, are used to evaluate the proposed scheme. Second, experiments are conducted on the ATM testbed of the High-Speed Networking and Performance (HSNP) Laboratory. An enhanced wireless ATM protocol is implemented to facilitate the desired mobility functions. Experimental data are collected through several experiments. Finally, experimental results based on the collected data are analyzed and presented.

To my parents and my wife

I

throu

edge

A. Sh

I a

ment

this

of ar

Evel

and

when

me a

ACKNOWLEDGMENTS

I would like to thank the continued support of my advisor, Herman D. Hughes, through the trials and tribulations of this Ph.D. Thesis. I would also like to acknowledge my thesis committee members, Vincent F. Melfi, Diane T. Rover, and Michael A. Shanblatt for their suggestions and advice.

I am indebted to my wife Pei-Yu for her support, understanding, and encouragement during my stay at Michigan State University. Without her love and support, this thesis would not have been possible. I am grateful to my parents for the gift of an education and the example of work ethic. A big thank you goes to Bryan and Evelyne for their help and friendship. Special thanks go to my brothers (Hui-Huang and Hui-Long) for their international long distance phone calls which propped me up when things got tough. I would also like to thank my daughter, Cheryl, for giving me a lot of joy and inspiration during the writing of this thesis.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES x

1 Introduction 1
- 1.1 Overview 1
- 1.2 Wireless ATM Networks 5
- 1.3 Motivation and Problem Statement 9
- 1.4 Objectives and Thesis Outline 12

2 Background and Related Work 17
- 2.1 Wireless Technologies 17
- 2.2 ATM Technology 20
- 2.3 PNNI Routing 25
- 2.4 Related Work on Location Management 27
  - 2.4.1 PCS Network Architecture Model 28
  - 2.4.2 IS-41 Location Tracking 30
  - 2.4.3 IS-41 Call Setup 33
  - 2.4.4 Improvements of IS-41 Location Management 34
- 2.5 Wireless ATM Projects 41
  - 2.5.1 Related Work on Handoff Management 47
- 2.6 Summary 50

3 Framework for WATM Location Management 53
- 3.1 WATM Network Architecture 53
- 3.2 WATM Location Management 57
  - 3.2.1 Location Tracking 58
  - 3.2.2 Call Setup 68
  - 3.2.3 Performance Analysis 73
- 3.3 Summary 79

4 Framework for WATM Handoff Management 81
- 4.1 Preliminary 81
- 4.2 WATM Handoff Management 84

4
4
4
4
5 E
5
5.
5.
5.
6 S
6.
6.
APP
A G
B P
B
B
B
B
B
B
B
B
B
BIBI

4.3 Intra-Domain Connection Handoff Management 89
4.3.1 Intra-Domain Uplink Connection Handoff 89
4.3.2 Intra-Domain Downlink Connection Handoff 96
4.4 Inter-Domain Connection Handoff Management 102
4.5 Preserving Cell Sequencing 110
4.6 Summary 115

5 Experiments **117**
5.1 Introduction 117
5.1.1 HSNP ATM Testbed 118
5.1.2 ATM on Linux Package 120
5.2 Experimental Setup 123
5.2.1 Architecture of VPWA Network Entities 125
5.2.2 Call Setup 130
5.2.3 Data Delivery 132
5.3 Experimental Results 135
5.4 Summary 140

6 Summary and Future Work **142**
6.1 Summary 142
6.2 Future Work 145

APPENDICES **148**

A Glossary **148**

B Performance Derivation **151**
B.1 Intra-Cluster Uplink With-Hint Analysis 151
B.2 Intra-Cluster Uplink Without-Hint Analysis 152
B.3 Intra-Cluster Downlink With-Hint Analysis 153
B.4 Intra-Cluster Downlink Without-Hint Analysis 155
B.5 Inter-Cluster Uplink With-Hint Analysis 156
B.6 Inter-Cluster Uplink Without-Hint Analysis 157
B.7 Inter-Cluster Downlink With-Hint Analysis 158
B.8 Inter-Cluster Downlink Without-Hint Analysis 160
B.9 Inter-Domain Phase 2 161
B.9.1 Uplink Case 161
B.9.2 Downlink Case 161

BIBLIOGRAPHY **163**

List of Tables

2.1 Typical application requirements 24

3.1 Definition of analytical parameters 74
3.2 Definition and expression of cost for various conditions 75
3.3 Signaling cost parameter sets 76

4.1 Augmented signaling message for supporting handoff 86
4.2 Definitions of parameters and notation 87

5.1 Experimental results in various parameter sets 136

4.14

5.1
5.2
5.3
5.4
5.5
5.6
5.7
5.8
5.9
5.10
5.11

4.14 Data and ACK formats for WATM sublayer 114

5.1 Configuration of HSNP ATM testbed 119
5.2 ATM on Linux protocol stack 121
5.3 ATM on Linux signaling procedure 121
5.4 Experiment setup: (a) the physical topology (b) the virtual topology . . 124
5.5 Architecture of the VPWA protocol stack with the MSU functions 127
5.6 Process structure of the network entity programs 128
5.7 Data structure of a connection control block at a BS 132
5.8 Trace of ATM cell stream with parameters of set 1 (every 100 msec) . . . 137
5.9 Trace of ATM cell stream with parameters of set 2 (every 100 msec) . . . 137
5.10 Trace of ATM cell stream with parameters of set 3 (every 100 msec) . . . 138
5.11 Trace of ATM cell stream with parameters of set 4 (every 10 msec) . . . 138

C

In

1.1

Wire

strat

years

emer

from

tions

tethe

tion

and

proce

tion

Chapter 1

Introduction

1.1 Overview

Wireless communications have evolved dramatically since Guglielmo Marconi demonstrated the first radio communication 100 years ago [1]. But during the past ten years, the wireless communication industries have grown by orders of magnitude. The emerging popularity of wireless communications may arise from the desire to be free from tethers, i.e., from physical connections to communication networks. Contributions to the surge of interest also include several other factors, such as: proliferation of tetherless personal computing, entertainment, and communication devices; liberalization of spectrum allocation procedures for both public personal communication system and private wireless local-area network (LAN) applications; advances in digital signal processing and radio modem technologies; improvements in cost/size/power consumption characteristics of digital electronics; etc. For example, wireless technologies for

two-
serv
mol
tele
com
any
[3. 4
appl
T
and
Nati
Term
comb
ment
and i
Euro
pean
Mobi
Telec
O
the p
troni

two-way messaging have been available for more than a decade [2]. Other wireless services, such as cordless telephone, cellular mobile radio systems, and satellite-based mobile systems, are now serving people's communication needs. However, current telecommunication networks still can not fulfill customers' dreams of *Star Trek*-style communications—receiving and transmitting voice, video, and data communication anytime, anywhere. But recent advances in Personal Communication Services (PCS) [3, 4, 5, 6, 7, 8] have prompted commercial vendors to explore many multimedia applications over an ubiquitous environment.

The idea of PCS is to provide diverse personal communication services globally and seamlessly. According to the most current definition of PCS in the American National Standards Institute (ANSI) approved standard, "Personal Communication Terminology" (ANSI T1.702-1995), PCS is "A set of capabilities that allows some combination of terminal mobility, personal mobility, and a service profile management." The development of PCS is expected to evolved from current wireless services and infrastructures to the global network, which is in the standardization process in Europe as the Universal Mobile Telecommunications System (UMTS) of the European Telecommunications Standards Institute (ETSI), and the Future Public Land Mobile Telecommunications System (FPLMTS) within the worldwide International Telecommunications Union (ITU) [9].

On the other hand, the wireline network has also gained significant attention over the past few years. Although Internet had been providing services, deploying electronic mail (E-mail), file transfer (FTP), telnet, and similar services around the world, the use of these services was usually confined to a relatively small population. New multimedia applications, such as World Wide Web (WWW), Internet phones, etc., have increased the public interest and contributed to the surge of Internet usage. Meanwhile, the increase usage exposed the weaknesses of the Internet technology: limited available bandwidth, no realtime transmission support, and no quality guarantee. Many Internet users currently use telephone lines and modems, which typically have very low transmission rates, to connect to the Internet. Hence, a large number of Internet links have limited transmission capability. However, thanks to the dramatic progress of hardware technologies in the areas of VLSI and fiber optics, the future wireline network will provide high-speed transmission from the range of Mbps to Gbps, which aims at providing immense transmission bandwidth and reducing communication delay.

In the mean time, the leaping progress of hardware has quickly made the communication protocol out of date. For example, Internet Protocol (IP) was initially designed for data communication and is not able to satisfy the level of Quality of Service (QoS) required for today's multimedia applications. In the future wireline backbone, known as Broadband-Integrated Services Digital Networks (B-ISDN) [10, 11, 12], multimedia applications are expected to dominate the network traffic. Equipped with a new suit of smart communication protocols, called Asynchronous Transfer Mode (ATM), B-ISDN promises to offer a wide range of multimedia applications including: data transfer, video conferencing, video-telephony, remote learning, and remote medical diagnostics. Also, networks such as the Public Switched Telephone Network (PSTN)

[1].

inte

I

tion

evol

ing

whic

appl

dem

imag

lieve

from

necti

N

Sinc

cons

[14,

and

prot

for n

envir

(WA

[1], the Internet, and Cable Television (CATV) will be replaced by B-ISDN, which integrates these services and applications all into one single network.

Even though the current PCS developments have concentrated on voice applications (cellular phone systems), it should be noted that these systems are inevitably evolving towards supporting a wide range of telecommunication applications involving packet data, video, and multimedia. Meanwhile, wireless local-area networks, which were initially designed for conventional data, are forced to support computer applications incorporating image and video transfer. Therefore, the parallel trends of demanding multimedia-capable services have sparked the integration of voice, video, image, and data in both telecommunication and computing environments. It is believed that the next generation of PCS will provide such multimedia services to users from wireline backbone networks. Therefore, the integration of mobile, wireless connections in a backbone broadband network is quite promising.

Now, the question is how to effectively integrate wireless and wireline networks. Since ATM [13] has been selected as the transport technology for the B-ISDN, it is considered advantageous to extend the ATM protocol to the wireless environment [14, 15]. Such extension will avoid the development of a new communication protocol and eliminate excessive overhead and delay due to protocol conversion. The extended protocol will also be able to provide certain levels of ATM-conformable QoS guarantee for multimedia traffic flowing between wireline and wireless networks. An integrated environment which employs the extended ATM protocol is called a **Wireless ATM (WATM)** network.

H

app

the

fixed

beca

chan

netw

F

the P

hand

netwo

or hu

QoS l

system

1.2

The ba

technol

environ

network

broadba

However, there are some issues that need to be resolved before seamless ATM applications across wireline and wireless networks can be supported. For example, the ATM protocol can not handle end user mobility since it is originally designed for fixed networks. Limited bandwidth in wireless networks also causes another problem, because a big pipe of traffic from a wired ATM network can easily overflow wireless channels. Therefore, extensive research must be done before the success of WATM networks can be achieved.

Figure 1.1 provides an overview of the future WATM network infrastructure for the PCS. By using the ATM network as the network backbone, mobile terminals or hand held devices will be able to access varied data and information on the ATM network. Legacy networks will connect to the ATM network via ATM multiplexors or hubs. Multimedia services, such as audio, video, and still images, with certain QoS levels will be retrieved by wired and wireless systems. Satellite communication systems will also be a part of the ubiquitous communication environment.

1.2 Wireless ATM Networks

The basic idea of a WATM network is to provide universal PCS by applying ATM technologies. The WATM protocol will extend the ATM protocol to the wireless environment with augmented functions to handle problems incurred in a wireless network. A WATM connection consists of paths (routes) going through the ATM broadband backbone network and wireless links between a Mobile Host (MH) and

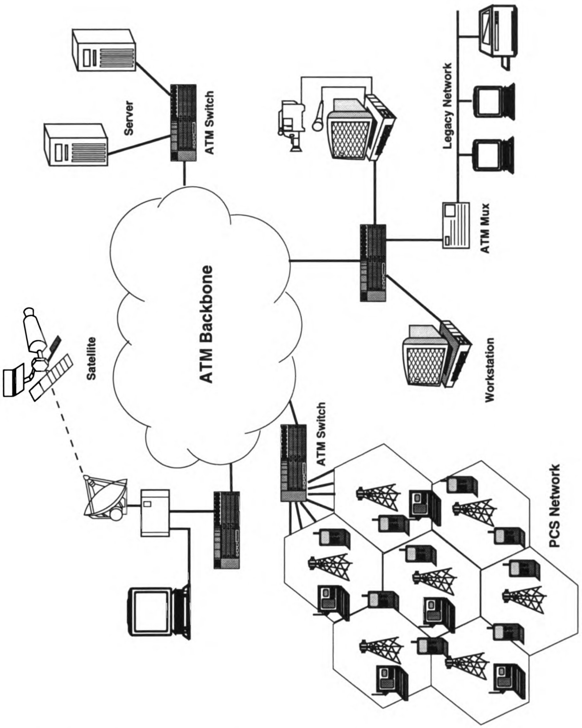


Figure 1.1: Overview of PCS communications in wireless ATM networks

a

ea

a

ap

dc

AT

S

wir

issu

a wired network access point, called Base Station (BS). Like current PCS systems, each BS in a WATM network will be responsible for wireless communications within a geographical region, called a *wireless cell.* BSs interconnect with each other by applying ATM switching as their backbone technology. A connection setup or tear-down between communication parties on a WATM network must be performed by an ATM-conformable signaling messages. The advantages of WATM include [14]:

- Compatibility with the B-ISDN/ATM base future wireline network,
- Flexible selection of service bit-rate,
- Statistical multiplexing efficiency for packet and bursty data terminals,
- Simplifying the interconnection problem of a large number of base stations associated with a micro/pico mechanism,
- Avoidance of protocol converting overhead.

Several challenging issues must be addressed before the success of integrating wireline and wireless networks, based on the ATM protocol, can be achieved. These issues are summarized as follow:

- *Bandwidth mismatch*: At present, ATM networks can provide 155 Mbps or 622 Mbps bandwidth on a single link. Giga-bit Network Interface Cards (NICs) will be soon available. Lower-speed NICs at 25/50 Mbps will also be available in the near future. Wireless networks, on the other hand, have relatively restricted bandwidth, around several Mbps depending on the physical layer technology used (spread spectrum or narrow band) and the frequency band allocated.

- *High error rate*: ATM was designed for such media as optical fibers whose bit error rates are extremely low (about 10^{-10}). However, the bit error rate of a wireless link is around 10^{-4} to 10^{-3}. The high error rate will seriously degrade the transmission efficiency of an ATM traffic stream. The other problem is that payload error within an ATM cell is not detectable because an ATM cell header only includes header error checking. Moreover, burst errors of wireless links will likely destroy routing information on the cell header because the simple header error checking mechanism is not able to recover the damaged header information.

- *User mobility*: The current ATM standard is based on fixed sources and destinations and does not take into account the movement of end systems. However, in the future WATM environment, a MH may move constantly from one location to another. To handle the user mobility problem in the WATM environment, the ATM connection establishment and routing mechanism must be modified. New signaling functions are also needed to track user location and to support connection rerouting. The ability to trace a MH's location and still maintain end-to-end virtual circuit connection poses a great challenge to the present ATM standard.

1.

The

solv

tech

spee

erro

ror (

[16.

are

mot

netw

loca

•

- *QoS guarantee*: A major advantage of the ATM protocol is that it can provide QoS guarantees depending on the contract with the application. ATM QoS algorithms are based on several assumptions: high speed transmission media, low error bit rate, and stationary end hosts. In the case of wireless network, however, these QoS guarantees are challenged because of high error bit rate, relative slow transmission rate and frequent end-host movements.

1.3 Motivation and Problem Statement

The aforementioned bandwidth mismatch problem in WATM network may be resolved by carefully assigning frequency bands and advancing the wireless hardware technologies. It is expected that 25 Mbps wireless modem, which matches the low speed of ATM standard, will be available soon. Researchers are investigating the high error bit rate issue by applying channel interleaving, ATM interleaving, Forward Error Correction (FEC), and data link Automatic Repeat ReQuest (ARQ) mechanisms [16, 17, 18, 19]. The QoS guarantee issue can be answered only if the other issues are solved. Among the many issues associated with WATM networks, handling user mobility is considered the most challenging one in regard to the success of WATM networks. In discussing the user mobility problem, we focus on two major issues: location management and handoff management [20].

- **Location management**

  Location management contains two functions: tracking the current position of a mobile and handling queries regarding the location of a mobile. When a call for a MH is initiated, the network must be able to locate the called MH and deliver the call. The current PCS network employs a two-level location management model from the IS-41 standard [21]. A MH must register to a home domain, called its *home location*. The *home location* is equipped with a location agent that keeps track of the current positions of all mobiles who have home registering in thaatt domain. When a mobile moves out of its *home location*, it must register itself to the *visiting location*. The agent in the visiting location will inform the mobile's *home location* of the mobile's current location. If a host tries to connect to the mobile, it must query the mobile's home agent to find out the mobile's current location. Upon obtaining the mobile's current location, the initiator can send a connection request to the *visiting location* following standard connection setup procedures. However, the location management mechanism of IS-41 incurs substantial overhead for querying a user location and delivering calls. Many improvements for this scheme have been reported in the literature. More detailed information on the IS-41 scheme and improvement scenarios is summarized in Chapter 2. Due to the expected high handoff frequency in the future WATM network, the location management mechanism of a WATM network must be able to update user location quickly and efficiently. On the other hand, unlike current PCS networks with two separate networks (PSTN and SS7), the WATM network is an integrated network, which will carry both user data and system signaling messages. Therefore, location management models based on the current PCS network are not suitable for emerging WATM networks. A more sophiscated location management scheme is needed for the WATM network.

- **Handoff management**

When a MH moves from one wireless cell to another while a connection is in progress, the continuity and the quality of the connection must be maintained. The backbone network must switch the access point from the previous BS to the BS which is currently serving the MH. The process of transferring the control of an on-going connection due to a MH's movement is known as the *handoff management*. Handoff management for WATM networks poses a great challenge to the current ATM protocol. ATM is a connection-oriented technology, with a connection establishment phase prior to data exchange. Once a connection is established, its routing path remains unchanged until data exchange is finished. However, in the WATM environment, end hosts are expected to move frequently from one location to another location. When the quality of a radio link between a wireless terminal and its BS degrades, a new BS with acceptable quality must be found, and network control functions of both the fixed and wireless network need to be invoked. In the backbone network, handoff requires the establishment of a new route, which transports the packets destined to (or originated from) the wireless terminal to (or from) the new BS. The handoff issue will be aggravated since the handoff frequency increases substantially when

1.4

The

user

mob

the future WATM geographical cell structure adopts either microcell or picocell architecture.

Current research on connection handoff mechanisms in WATM has focused on providing algorithms for preserving connections of mobile end hosts [22, 23, 24, 25, 26, 27]. When a MH moves to another wireless cell, queued and incoming ATM cells at the old BS can not be delivered to their destination as the BS is no longer communicating with the MH. An intuitive solution is to discard these ATM cells and let the higher protocol layers at the end systems handle the problem [24]. In this case, lost cells result in damaged packets at the higher protocol layer, such as Service Specific Connection Oriented Protocol (SSCOP) or Transmission Control Protocol (TCP). For loss sensitive applications, end-to-end retransmissions are issued to retrieve those damaged packets. For realtime applications, retransmissions of packets are not practical since they reach the destination too late. In such a case, the QoS is degraded due to lost cells. It is clear that an efficient connection handoff scheme is needed to resolve these problems.

1.4 Objectives and Thesis Outline

The main objective of this research is to provide a schematic scenario for handling user mobility in future WATM networks. The challenging issues of supporting user mobility in WATM networks to be addressed by this research include:

- *Designing WATM network architecture for PCS services.* A WATM network architecture is desirable for migrating PCS services to ATM-based backbone networks.

- *Facilitating user location management functions.* In order to locate mobile users and to deliver calls, a location management scheme, similar to that of the current PCS network, is needed.

- *Maintaining end-to-end connections when the end host constantly changes its location.* Since ATM is a connection-oriented technology, an established connection must be preserved even when MHs movements are involved.

- *Preserving the cell sequence of a handoff connection.* This is essential because out-of-sequence cells will cause the data frames at the higher protocol layer to be discarded.

- *Minimizing the connection disruption time at the connection handoff.* For real-time applications, the connection disruption time will result in the degradation of QoS.

- *Reducing the cell loss probability due to connection handoffs.* Cell loss usually degrades the QoS of realtime application or causes the retransmission of packets for non-realtime traffic.

- *Fitting the handoff-supported functions into the ATM protocol standard.* Augmented handoff functions must be incorporated with the current ATM protocol af

toc

illu

ica

AT

me

link

pro

toc

sui

on

the

dlin

Di

for

tec

ing

use

standard in order to avoid the overhead of the protocol conversion and the development of a new protocol.

The primary focus of this research is to develop and analyze a protocol with the aforementioned features. With regard to the current ATM protocol stack, this protocol will impact both the control signaling layer and the ATM layer. Figure 1.2 illustrates the layer structure of a WATM protocol, which adds a new wireless physical layer, medium access control sub-layer, and data-link sub-layer to the standard ATM protocol stack. Most of mobility support functions (including location management and handoff management) reside in the control signal layer. The WATM data link sublayer and media access control sublayer are implemented mainly to deal with problems caused by wireless links and provide media transparency at the higher protocol layer. In the design of our handoff management scheme, the WATM data link sublayer is involved in preserving cell sequencing of a handoff call. Further discussion on this design is presented in section 4.5 of Chapter 4.

The remainder of this thesis is organized as follows. In Chapter 2, we provide the background information regarding WATM networks and related work on handling user mobility in current wireless systems. Wireless technologies, such as Time-Division Multiple Access and Code-Division Multiple Access, have different strategies for wireless transmission of data. The advantages and disadvantages of these wireless technologies are discussed. General concepts of ATM technologies, needed in depicting future WATM networks, are also briefly summarized. Related work in handling user mobility (including location management and handoff management) is surveyed.

Th

sea

on-

har

arc

Th

me

tec

loc

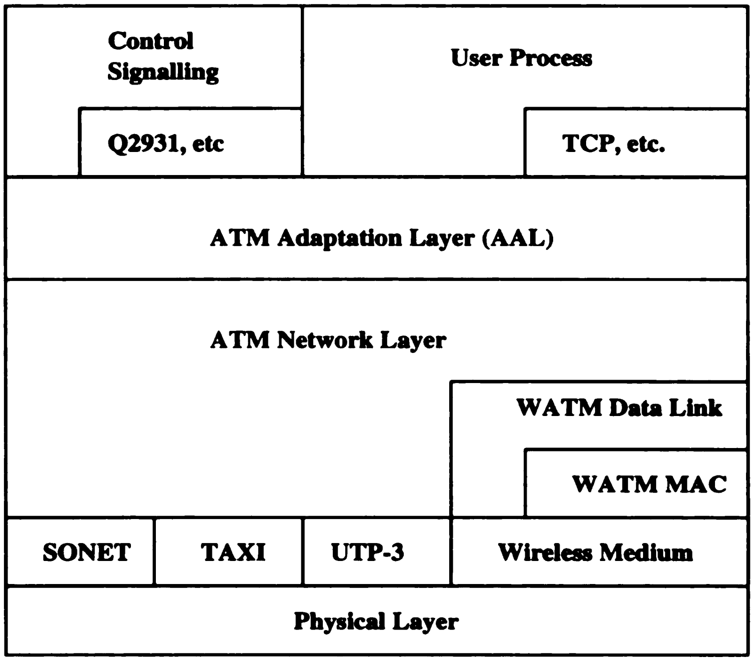


Figure 1.2: Wireless ATM protocol layers

The current IS-41 location management model is presented first. Then current research in improving the IS-41 location management is introduced. We also describe on-going WATM projects in both universities and industries, with emphasis on the handoff management strategies used.

In Chapter 3, first a new WATM network architecture is introduced. This network architecture separates the core ATM backbone network from the peripheral network. This separation allows us to provide user mobility support functions without major modification of the deployed ATM backbone network. Based on this network architecture, we propose a new location management scheme for effectively tracking user location and setting up calls in future WATM networks.

A handoff management mechanism is the other major task needed to support user mobility in a WATM network. Chapter 4 describes the handoff management mechanism we proposed. This handoff management scheme makes use of the previously reported WATM network architecture to support fast and efficient connection handoffs.

To help examine the proposed user mobility management framework, we conduct experiments on the ATM testbed of High-Speed Networking and Performance (HSNP) Laboratory. Chapter 5 describes the experimental setups and the associated performance results. In Chapter 6, we summarize the main contributions of the research and discuss some possible directions for future work.

Chapter 2

Background and Related Work

This chapter presents background information related to the study of handling user mobility in WATM networks. Wireless technologies, ATM network technologies, and issues of WATM networks are covered. In addition, current research relative to both location management and handoff management is surveyed.

2.1 Wireless Technologies

Today's wireless LAN technologies comprise infrared (IR), UHF radio, spread spectrum, and microwave radio, with frequencies in the region of 900 MHz in the U.S. (GHz in Europe) to infrared frequencies [28]. The Personal Communication Network (PCN) may use a shared wideband Code-Division Multiple Access (CDMA) band, or Time-Division Multiple Access (TDMA). There is a considerable debate among researchers regarding the relative merits of spread spectrum (CDMA), and narrow-ba

th

band (TDMA) for PCNs [29, 15, 20]. The preferred technique may actually vary with the specific PCN application scenario.

- *Spread Spectrum (CDMA)*: The term spread spectrum defines a class of digital radio systems in which the occupied bandwidth is considerably greater than the information rate [28]. The technique was initially proposed for military use, where the high levels of difficulty of detecting or jamming such a signal made it an attractive choice for covert communication. CDMA is often used in reference to spread spectrum systems and refers to the possibility of transmitting several such signals in the same portion of spectrum by using pseudorandom codes for each one. This can be achieved by either frequency hopping (a series of pulses of carrier at different frequencies in a predetermined pattern) or direct sequence (a pseudorandom modulating binary waveform symbol rate is a large multiple of the bit rate of original bit stream).

- *Time Division Multiple Access (TDMA)*: The basic principle of TDMA is simple. Traditionally, voice channels have been created by dividing the radio spectrum into (very narrow frequency) radio frequency (RF) carriers (channels), with one conversation occupying one (duplex) channel [28]. This technique is known as Frequency Division Multiple Access (FDMA). TDMA divides the radio carriers into an endlessly repeated sequence of small time slots (channels). Each conversation occupies just one of these time slots. Instead of just one conversation, each radio carrier can support a number of conversations at once.

m

in

is

CI

on

sw

Th

dy

one

for

sign

One of the most fascinating properties of CDMA is that it is capable of supporting multiple simultaneous users on the same wideband channel. It provides an efficient integrated solution for frequency reuse and multiple access. Another merit of CDMA is that, unlike FDMA or TDMA, there is no "hard" capacity limit. However, a CDMA receiver requires that the signal strengths of each received signal be equal. If one transmitter is significantly closer to the receiver than all others, its signal may swamp the receiver's front-end and prevent the reception of the other signals [30]. This is the *near-far problem*, which adds to the system complexity by requiring a dynamic power control scheme to resolve this problem.

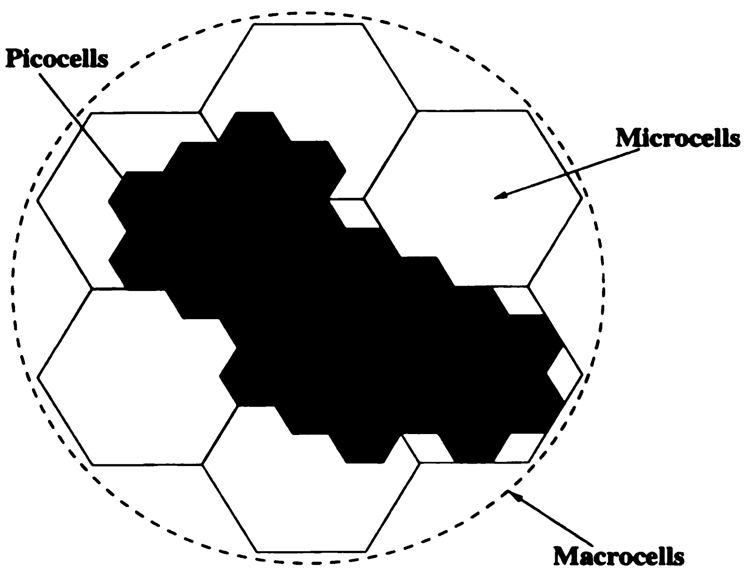


Figure 2.1: Geographical structure of wireless cells

In wireless networks, a mobile terminal communicates with a control point (BS) via one of the aforementioned wireless technologies. A BS serves as a control entry point for MHs to connect with a wired backbone network, such as the PSTN. As a wireless signal decays according to the distance between the transmitter and the receiver [1], a BS can only communicate with MHs within a certain distance, depending on the wireless technologies used. The geographical area covered by a BS's wireless signal is called a *wireless cell*. Usually the coverages of neighboring wireless cells will have an overlap area, called a *handoff zone*. Depending on the wireless technologies, the cell size can be categorized into: macrocell ($\sim$ several miles), microcell (1000–2000 feet), and picocell (100–200 feet), as shown in Figure 2.1. In case of the TDMA scheme, neighboring cells cannot use the same frequency. In order to prevent interference, frequency assignment needs to be carefully planned in advance. This is generally known as the coloring problem. But for CDMA, the whole spectrum of frequency is used on every cell.

2.2 ATM Technology

Asynchronous Transfer Mode (ATM) is standardized as a transport vehicle of the backbone B-ISDN network. ANSI and ITU have defined ATM as the standard for the transport of a complete range of user traffic, including voice, data, and video signals, on any User-to-Network Interface (UNI) [13]. ATM is a packet-switching, connection-oriented, and point-to-point connection networking technology. It is a technology based on 53-byte packet units, called *cells*, for transporting user information within the Synchronous Optical Network (SONET) payload [31]. Every ATM end host is identified by a 20-byte long ATM address which is defined in the ATM Forum standards. The 20-byte ATM address is based on the Network Service Access Point

(N
4S
as
sw
sp

to-
VC
tra
(C.
acc
dec
the
for
the

tin
tio
(VI
is c
pa
the

(NSAP) addressing format [13]. Each ATM cell is comprised of 5 bytes of header and 48 bytes of data payload as shown in Figure 2.2. The fixed format of the two fields assures that the processing may be done by simple hardware circuits, called ATM switches [32, 33, 34, 35, 36, 37, 38, 39]. Thus, ATM switches enable faster processing speeds than do software implementations [40].

Before end entities on an ATM network can communicate with each other, an end-to-end connection, called Virtual Channel (VC), must be established first. To setup a VC connection, a caller sends a connection request to the ATM network with a set of traffic descriptors which represent the QoS of this connection. Call Admission Control (CAC) [41, 42, 43] of ATM networks processes the request and decides whether to accept the connection request or not, based on the current network status. This decision to accept is made in a way not to affect the QoSs of current connections. After the connection request is accepted, buffer and transmission bandwidth are reserved for this connection along its routing path. Then the end systems can proceed with their communications.

During communications, data generated by various applications at the transmitting side have been segmented and encapsulated into ATM cells at the ATM Adaptation Layer (AAL). Cells belonging to a channel are identified by Virtual Path Identifier (VPI) and Virtual Channel Identifier (VCI) in their headers [13]. The identification is only meaningful locally and may be changed at each switch point. The VPI/VCI pair is assigned by ATM networks at the connection setup phase. By examining the VPI/VCI information in headers of ATM cells, ATM switches deliver these cells

tc
as
ce
ca
(V)
a s
stra

toward their destinations. All cells associated with a given VC flow over the route assigned to that connection and are delivered in sequence. At the receiving end, ATM cells are reassembled into complete messages via the AAL layer. Note that messages can be either an individual packet or a continuous bit stream.

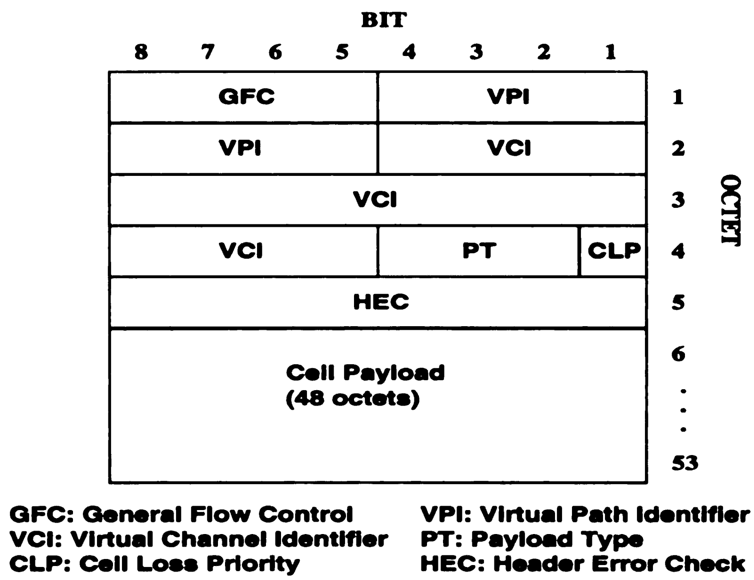


Figure 2.2: ATM UNI cell format

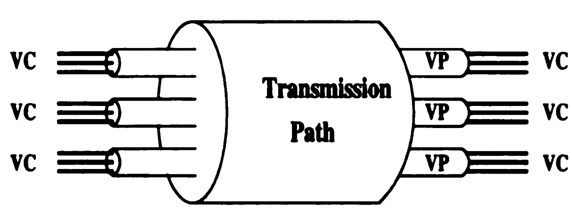


Figure 2.3: Virtual path and virtual channel structure

ATM VC connections with the same VPI traverse along the same Virtual Path (VP). The concept of a VP is introduced to provide the capability for manipulating a set of ATM connections as one unit. The structure is shown on Figure 2.3. This strategy allows ATM networks to avoid wasting unused bandwidth in circuit switching te

ca

st

ty

tic

Ta

isf

cor

me

wh

rat

tin

tio

of

ab

CB

an

era

M

[4

technology due to low traffic load on the dedicated circuit. Multiple traffic streams can be multiplexed asynchronously and delivered altogether in a VP.

Multimedia applications are known for their diverse traffic patterns, time constraints, and allowable data loss rate. For example, realtime audio traffic, which is typically sensitive to end-to-end delay and delay jitter, can tolerate certain information loss. File transfer will not tolerate data loss but have more loose time constraints. Table 2.1 shows traffic requirements of some typical multimedia applications. To satisfy varied multimedia traffic, the ATM protocol is equipped with a suite of smart control mechanisms, which promise to handle diverse traffic requirements [41]. As mentioned earlier, a set of traffic descriptors must be submitted to the ATM network when issuing a connection request. The set of traffic descriptors includes peak cell rate (PCR), sustained cell rate (SCR), mean cell rate (MCR), and cell delay various time (CDVT). These parameters are used to describe the traffic pattern of a connection and to calculate the demanded bandwidth and buffer size by the CAC module of ATM networks.

The ATM Forum has classified ATM traffic into: Constant Bit Rate (CBR), Variable Bit Rate (VBR), Available Bit Rate (ABR), and Unspecified Bit Rate (UBR). CBR traffic requires periodical transmission with certain bandwidth. Telephone calls and uncompressed broadcast video fit into the CBR category. VBR applications generate traffic on peak bit rate at one time and may later change to lower bit rate. Motion Picture Expert Group (MPEG) video applications is the typical example [44, 45, 46, 47]. VBR includes both realtime and non-realtime service classes (VBR-

Applications	Service Type	QoS	Bit-rate range
Voice telephony	CO/CBR	Call blocking permitted; Low-med cell loss OK; Isochronous	2.4–32 Kbps
Digital audio	CO/CBR	Call blocking permitted; Low cell loss; Low delay jitter	128–512 Kbps
Teleconference, Multimedia Comm., Digital video	CO/CBR or CO/VBR	Statistical mux (for VBR); Call blocking permitted; Low-med cell loss OK; Low delay jitter	15–20 Mbps
Digital HDTV	CO/CBR	Call blocking permitted; Low-med cell loss OK; Low delay jitter	15–20 Mbps
General computer data	CL, Best effort	No call blocking; Low cell loss; Med delay &jitter OK	0.1–1 Mbps
E-mail	CL, Best effort	Low transfer rate; No call blocking; low cell loss OK; High delay OK	9.6–128 Kbps
High-speed data (file transfer, multimedia)	CL, Burst mode	High transfer rate; Very low cell loss; Med delay & jitter OK	1–10 Mbps

Table 2.1: Typical application requirements

RT

er.

tc

ge

pi

co

sc

tw

Bu

cel

occ

inf

2.

Pr

fir

m

in

ar

ne

si

RT and VBR-NRT). CBR and VBR-RT traffic usually has strict time constraint on end-to-end delay and cell jitter. Cells which cannot reach their destination according to these time constraints are considered out-of-date and lost. ABR traffic, however, generally tolerates longer transmission delay. Additionally, this type of traffic explores the network available bandwidth by a special flow control mechanism. The control scheme used by ABR traffic helps guard against data loss by instructing the source to throttle back when the network is congested. FTP and remote login are two examples. UBR traffic, like ABR, makes use of available network bandwidth. But there is no guarantee that data will arrive at its destination. This type of data cells are marked with low priority on their cell headers. When network congestion occurs, these cells are dropped first, since UBR traffic usually carries less important information.

2.3 PNNI Routing

Private Network-to-Network Interface (PNNI) is the standard routing protocol defined by the ATM Forum [48]. It is a topology-state algorithm that advertises information about the status of links and nodes in the network. In addition to topological information, the algorithm advertises address reachability. PNNI supports a hierarchical organization of the topology database. Using PNNI, switches in an ATM network automatically form a hierarchy of peer groups according to addresses assigned by the network manager. The switches' ATM addresses provide the key to the

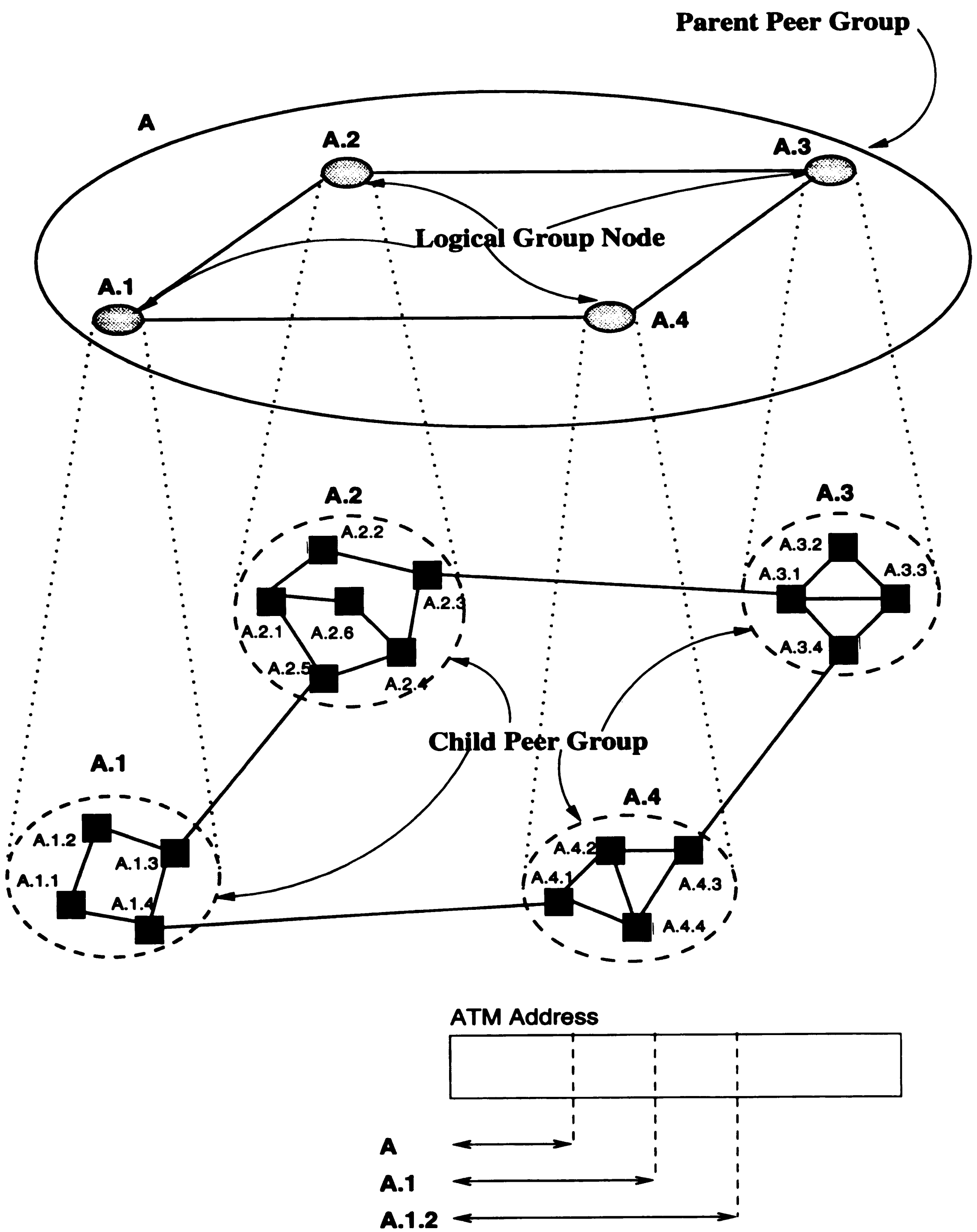


Figure 2.4: PNNI hierarchy structure.

structur

telepho

of the g

hierarch

tion for

the lowe

organiza

Once

informat

Circuits

the way i

a way to

metrics a

straint. F

path sele

istrative

maximur

2.4 I

To handl

mechanis

structure of this hierarchy. Each peer group has its own address identifier, similar to a telephone exchange or area code. A node is elected by the peer group to be the leader of the group. The leader performs additional functionality at the higher level of the hierarchy. An example of this functionality is the aggregation of topological information for presentation to the higher levels and advertisement of a high-level topology to the lower hierarchical levels. Figure 2.3 shows an example of a hierarchical topology organization.

Once the PNNI hierarchy is created, peer group leaders are elected, and routing information is exchanged, the ATM switches can begin to establish Switched Virtual Circuits (SVCs) between various end stations on the network. Because PNNI defines the way in which SVCs between ATM switches are established, it also must provide a way to maintain QoS guarantees. The PNNI routing decision is based on multiple metrics and attributes. If an attribute value for a parameter violates the QoS constraint, PNNI excludes that topological element from consideration while making a path selection. The metrics and attributes used in evaluating a route include administrative cost, maximum cell transfer delay, cell delay variation, cell loss ration, and maximum cell rate.

2.4 Related Work on Location Management

To handle the mobility of terminals in a WATM network, a location management mechanism is needed to track the position of a mobile (location tracking) and handle queri

tems. th

tronic T

(IS-41) [

cation pa

Mobile F

System (

ital Cell

function

manager

refer to [

2.4.1

Figure 2

stations

BSC is t

mobile s

dinates

wireline

tem No.

the SS7

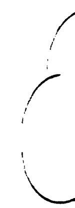
setup ar

dle queries regarding the location of a mobile (call setup). In today's PCS systems, there are two standards for location management currently available: Electronic/Telecommunications Industry Associations (EIA/TIA) Interim Standard 41 (IS-41) [21] and the Global System for Mobile Communications (GSM) mobile application part (MAP) [49]. The IS-41 is the standard of North America for Advanced Mobile Phone System (AMPS), IS-54, IS-136, and Personal Access Communication System (PACS) networks. GSM MAP is commonly used in Europe for GSM and Digital Cellular Service at 1800 MHz (DCS-1800) networks. The location management functions of both standards are very similar. To avoid redundancy, only the location management of IS-41 is presented in the following sections. Interested readers may refer to [49] for detailed information of the GSM MAP mobility management scenario.

2.4.1 PCS Network Architecture Model

Figure 2.5 shows the general network architecture of a PCS network. Several base stations are connected to a base station controller (BSC). The main function of the BSC is to manage the radio resources of its base stations. The BSC is connected to a mobile switching center (MSC), which provides typical switching functions and coordinates location tracking and call setup. The MSC is connected to both the backbone wireline network (such as PSTN) and the signaling network (e.g., the Signaling System No. 7, or SS7 network). The wireline network carries actual user calls, whereas the SS7 signaling network performs the network management functions, such as call setup and location tracking. The SS7 network [50, 51, 52], as shown in Figure 2.5, conne

messa



The

two typ

register

connects the HLR, the VLR, and the MSCs in a PCS network. The delivery of signal messages is managed by the Signal Transfer Points (STPs).

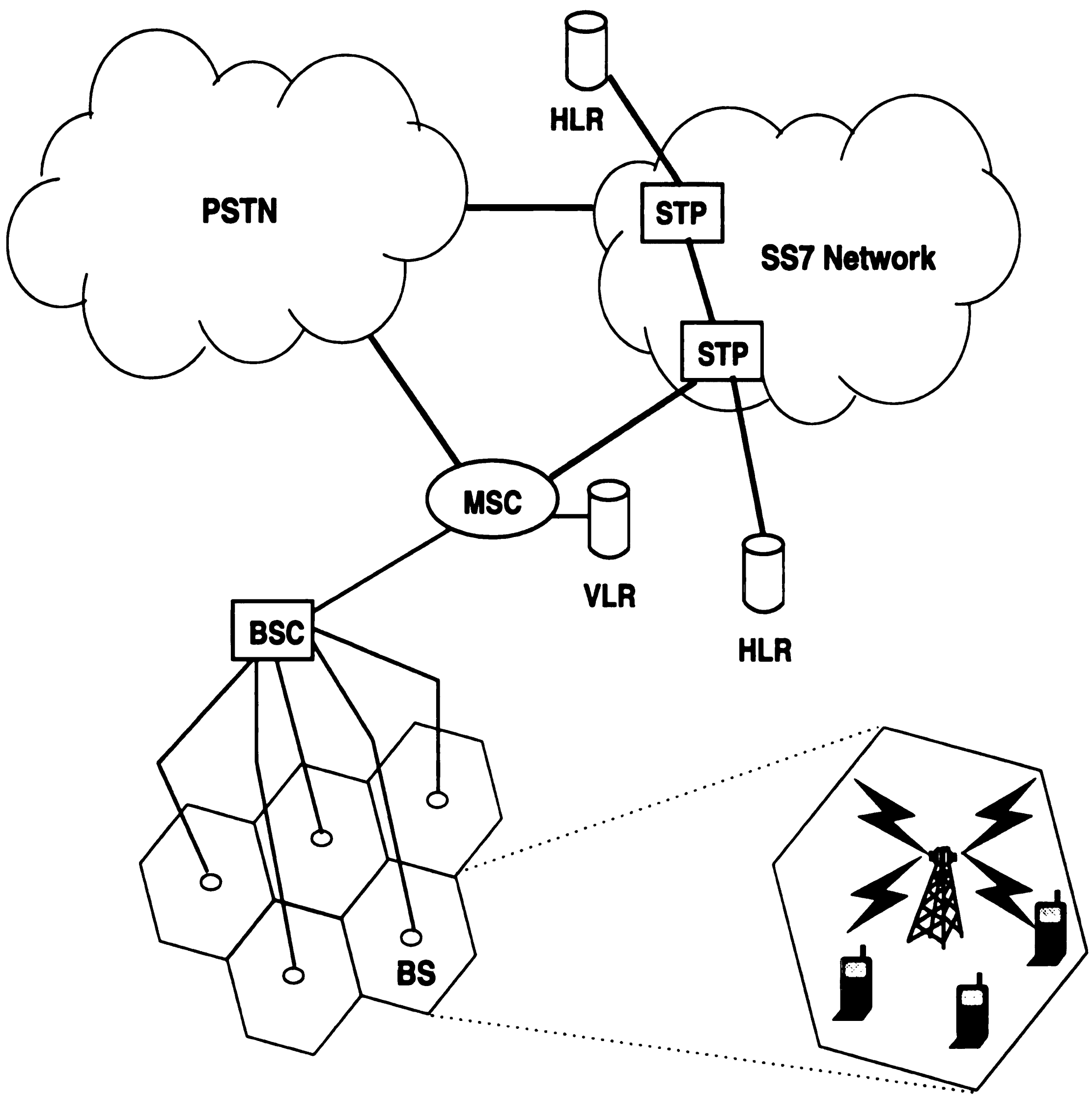


Figure 2.5: PCS Network Architecture Model

The IS-41 location management is based on a two-level data hierarchy such that two types of database, the home location register (HLR) and the visitor location register (VLR), are involved in the tracking of a MH [53]. A number of wireless cells form a

connect

Gen

HLR in

of servi

in a use

(retriev

2.4.2

When a

in these

success

informa

Loc

tion are

to the

record

belongs

- R

- U

- C

form a registration area. All base stations belonging to the same registration area are connected to the same MSC.

Generally, each PCS network has a HLR. A user is permanently associated with an HLR in the subscribed PCS network. Information about each user, such as the types of services subscribed, billing information, and location information, are maintained in a user profile located at the HLR. Each VLR keeps the information of the MHs (retrieved from the HLR), which are visiting its associated area.

2.4.2 IS-41 Location Tracking

When a MH moves around the network coverage area, the location information stored in these databases may no longer be accurate. To ensure that calls can be delivered successfully, a mechanism is needed to update the databases with up-to-date location information. This process is called **location tracking**.

Location registration is initiated by a MH whenever it moves into a new registration area. When a MH enters an registration area, if the new registration area belongs to the same VLR as the old registration area, the record at the VLR is updated to record the ID of the new registration area. Otherwise, if the new registration area belongs to a different VLR, the following operations must be performed:

- Register the MH at the new serving VLR
- Update the HLR to record the ID of the new serving VLR
- Cancel the MH at the old serving VLR

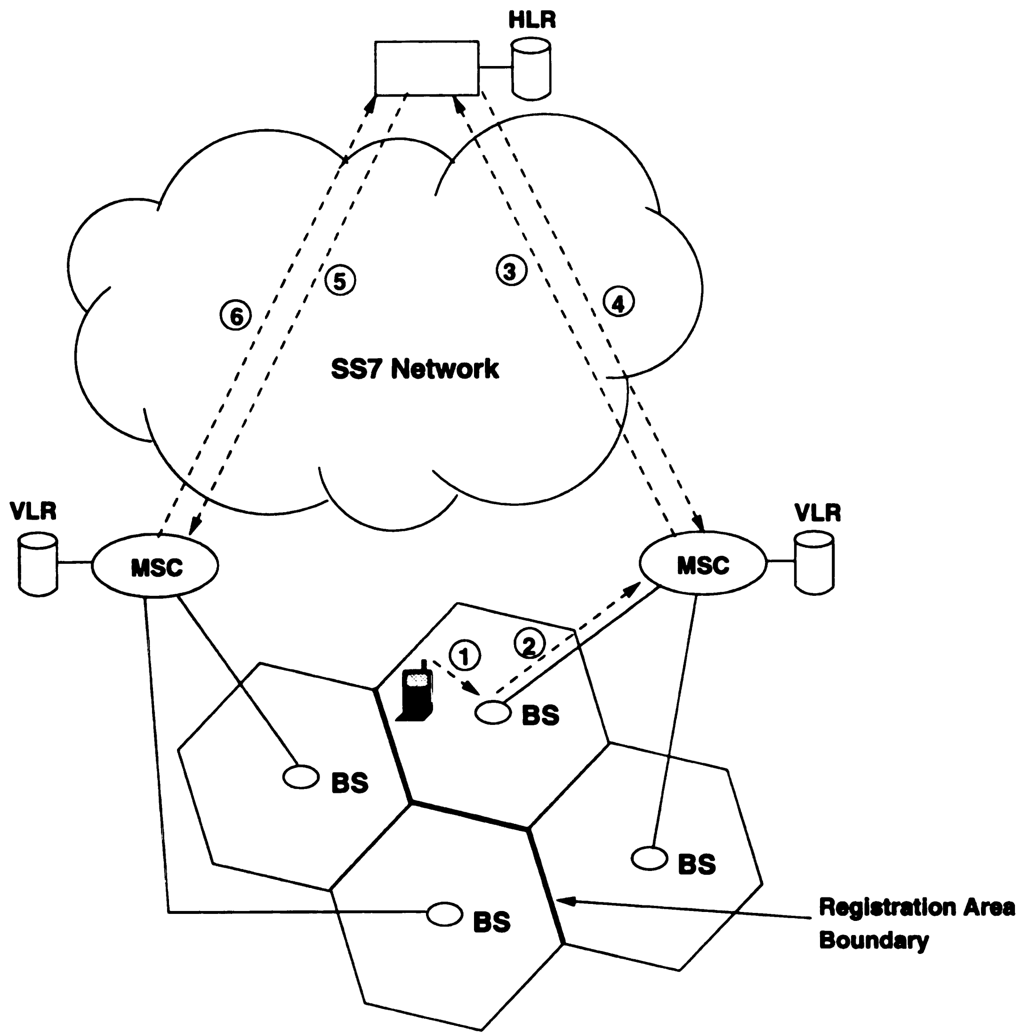


Figure 2.6: IS-41 Location Tracking Model

Fig

registr

locatio

1. C

r

2. T

w

3. T

a

o

l

l

c

4. T

th

a

5. T

6. T

e

Figure 2.6 shows the location tracking procedure when a MH moves to a new registration area. The following is the ordered message flow that is performed during location tracking [53, 54]:

1. Once a MH enters a new registration area, the MH sends a location update request to the new base station.

2. The base station forwards the location update message to the MSC through a wired network, which launches a registration query to its associated VLR.

3. The VLR updates its record on the location of the MH. If the new registration area belongs to a different VLR, the new VLR determines the address of HLR of the MH from its mobile identification number. This is achieved by a table lookup procedure called global title translation. The new VLR then sends a location registration message to the HLR; otherwise, location registration is complete.

4. The HLR performs the required procedure to authenticate the MH and records the ID of the new serving VLR of the MH. The HLR then sends a registration acknowledgment message the new VLR.

5. The HLR sends a registration cancellation message to the old VLR.

6. The old VLR removes the record of the MT and returns a cancellation acknowledgment message to the HLR.

2.4.3

When a

informa

diagran

the act

1. T

th

2. T

g

th

3. T

r

s

4. T

b

T

5. T

t

6. T

n

2.4.3 IS-41 Call Setup

When a call is initiated for a MH, the call is delivered as a function of the location information acquired from the location tracking scheme. The signaling message flow diagram for the IS-41 call setup is shown in Figure 2.7. The following steps describe the activities that take place during call setup in [53, 54].

1. The calling MH sends a call initiation signal to the serving MSC of the MH through a nearby base station.

2. The MSC determines the associated HLR of the called MH by performing a global title translation procedure and then sends a location request message to the HLR.

3. The HLR determines the serving VLR of the called MH and sends a route request message to the VLR. This VLR then forwards the message to the MSC serving the MH.

4. The MSC assigns a temporary identifier called a temporary local directory number (TLDN) to the MH and returns a response to the HLR together with the TLDN.

5. The HLR forwards this information to the MSC of the calling MH in response to its location request message.

6. The calling MSC requests a call setup to the called MSC through the SS7 network. This completes a call setup.

Th

the ca

cell is

curren

proced

registra

MH se

cell.

2.4.4

Queryi

for call

current

is trans

area th

from a

in Seat

caller.

since th

by loca

for red

59, 60,

The procedure described above allows the network to set up a connection from the calling MH to the serving MSC of the called MH. Because more than one wireless cell is associated with a registration area, a paging (or alerting) procedure is used in current PCS networks to determine the cell location of the called MH. The paging procedure is achieved by broadcasting polling signals to all cells within the residing registration area of the called MH. Upon receiving the polling message, the target MH sends a reply which allows the MSC to determine its current residing wireless cell.

2.4.4 Improvements of IS-41 Location Management

Querying of location databases to determine the current location of a called MH for call setup may incur substantial overhead. For example, if the calling MH is currently located in Miami and its HLR is in Seattle, a location registration message is transmitted from Miami to Seattle whenever the MH moves to a new registration area that belongs to a different VLR. Hence, when a call for the MH is originated from a calling MH in Miami, the MSC of the calling MH must first query the HLR in Seattle before it finds out that the called MH is located in the same area as the caller. Although this may not be a significant problem in today's cellular network since the penetration is relatively low, the increase of the signaling traffic generated by location management is becoming extremely high [55, 56]. A number of schemes for reducing signaling traffic caused by IS-41 are available in the literature [57, 58, 59, 60, 61, 62, 63]. Depending on the organization of databases, these proposed


VLR

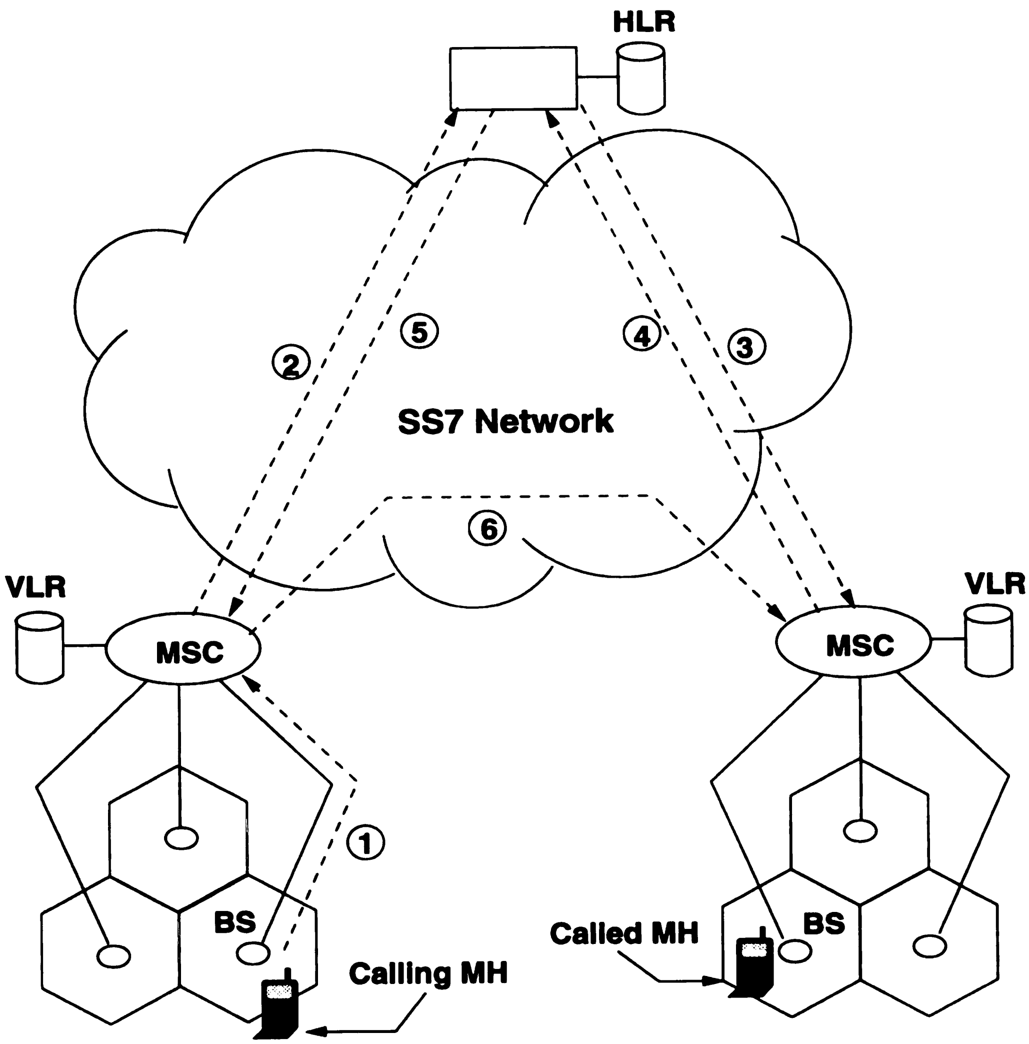


Figure 2.7: IS-41 Call Setup Procedure

mechan

and dist

IS-41 lo

while pr

of resea

scheme

are brie

2.4.4.1

Per-us

In [57],

scheme

a MH

signal

Whene

contain

is initia

cache e

is still r

has alre

miss oc

Con

mechanisms can be categorized into two groups: centralized database architecture and distributed database architecture. The first group is based on an extension of the IS-41 location management strategy. These schemes aim to improve the IS-41 scheme while preserving the centralized database architecture from IS-41. The second group of research mechanisms applies new database architectures that require a new set of schemes for location tracking and call setup. These mechanisms from both categories are briefly presented in the following subsections.

2.4.4.1 Centralized Database Architectures

Per-user Location Caching

In [57], the per-user location caching scheme is proposed. The basic idea of this scheme is that the volume of the signaling and database access traffic for locating a MH can be reduced by maintaining a cache of location information at a nearby signal transfer point (STP), which is responsible for routing the signaling message. Whenever the MH is accessed through the STP, an entry is added to the cache which contains a mapping from the ID of the MH to that of its serving VLR. If another call is initiated for a MH, the STP first checks if a cache entry exists for the MH. If a cache entry exists, the STP will query the VLR as specified by the cache. If the MH is still residing under the same VLR, a hit occurs, and the MH is found. If the MH has already moved to another location which is not associated with the same VLR, a miss occurs, and the IS-41 call setup scheme is invoked to locate the MH.

Compared to the IS-41 scheme, per-user location caching allows the STP to locate the VLR of the called MH with only one cache lookup. This is true, however, only when the cached location information of the called MH is valid (a hit). The cost of per-user location caching is higher than the IS-41 scheme when a miss occurs. Based on the system parameters, the minimum hit ratio required to produce a performance gain using per-user location caching has yet to be determined.

User Profile Replication

As described in [58], user profiles are replicated at selected local databases. Once a call is generated for a remote MH, the local database is checked first to determine if replication of the called MH's user profile is available or not. If the user profile is found, no HLR query is needed. The called MH can be located via the location information available at the local database. Otherwise, the IS-41 call setup is used to locate the called MH. Whenever the MH moves to another location, its current location information must be updated at all replications of the MH's user profile.

The advantage of this scheme is that the signaling and database access overhead for local management can be reduced significantly, depending on the mobility rate of the MH and the call arrival rate from each location. However, this scheme requires the central system to collect the mobility and calling parameters of the entire user population in order to make a replication decision. This may not be feasible because it requires data retrieval from a large number of PCS network providers. Hence, performing replication decisions and distributing replicated user profiles are computation-intensive and time-consuming processes which may demand a significant amount of network resources.

Pointer Forwarding

A pointer forwarding scenario is presented in [59]. Pointer forwarding eliminates the need for reporting location change of a MH to its HLR every time the MH moves to an area belonging to a different VLR. A forwarding pointer, pointing to the new VLR, is maintained at the old VLR. Thus, when a call for the MH is initiated, the network locates the MH by first determining the VLR at the beginning of the pointer chain and then follows the pointers to the current serving VLR. To avoid excessive searching time, the length of the pointer chain is limited to a pre-defined maximum value, K. If the length of the pointer chain reaches K, the pointer chain can not extend further. The next movement of the MH has to be reported to the HLR. The original pointers are deleted and the pointer chain is reset.

The performance of this scheme is highly dependent on the mobility and call arrival parameters and the maximum length of a pointer chain. It can not guarantee a reduction in cost from the original IS-41 scheme in all cases.

Local Anchoring

In [60], a local anchoring mechanism is introduced to reduce significant signaling traffic of location tracking by eliminating the need to report location changes to the HLR. Location information is sent to a VLR close to the MH instead of delivering this information to the HLR. The selected VLR acts as a local anchor for all location queries of the MH. The HLR only keeps a pointer to the local anchor. Because the local anchor is close to the MH, the signaling cost incurred in location tracking is reduced. The local anchoring has two approaches: static local anchoring and dynamic local anchoring. In static anchoring scheme, the serving VLR of a MH during its last call arrival becomes its local anchor. Although static local anchoring completely eliminates the need to report location changes to the HLR, it may not always result in performance improvement, the same problem encountered by the pointer forwarding and location caching schemes. On the other hand, the dynamic local anchoring changes the local anchor to the serving VLR when a call arrives. The network can also decide if a new anchor should be chosen according to the mobility and call arrival parameters. The authors (Ho and Akyikliz) show that this scheme is able to reduce the overhead of location tracking in any case.

2.4.4.2 Distributed Database Architectures

Tree-Base Distributed Database Approach A

In [61], a distributed database architecture for location tracking is introduced. Instead of applying the two-level HLR/VLR database architecture of IS-41, a large number of location databases are used to keep track of the current location of MHs. The database of this scheme has a tree structure. All MHs can associate with leaf nodes only. Every leaf node is equipped with a location database which contains location information of MHs residing in its subtree. To locate the called MH of a generated call, the tree hierarchy is traversed by following database entries. When a MH moves to a new area covered by another leaf node, the corresponding databases are updated to indicate the correct location of the MH.

Compared to centralized database schemes, this technique reduces the distance traveled by signaling messages. However, the advantage comes with the cost of increasing the delay in location tracking and call setup.

Tree-Base Distributed Database Approach B

Another distributed database mechanism is proposed in [62]. This scheme is very similar to the previous one except that MHs can associate with any node of the tree hierarchy. The root of the tree must have a location database. All other nodes on the tree may or may not have location databases. These databases contain pointers to MHs. If a MH is residing at the subtree of a database, a pointer is set up in this database pointing toward the residing node of the MH by following a pointer chain. Therefore, when a call for a MH is generated at any node on the tree, the called MH can be located by following these pointers of the MH. Although this scheme may reduce the number of database accesses and updates, it requires a special arrangement of these databases.

Partitioning Scheme

A partitioning scheme is introduced in [63]. This scheme groups location servers into partitions based on the fact that the mobility patterns of MHs vary among locations. The location tracking procedure is invoked only when a MH moves to a new partition.

Location registration is performed if a MH moves within a partition. In that case, the number of location registrations can be minimized in the area where the mobility rate of MHs is high. However, the cost reduction is highly dependent on the mobility and call arrival patterns as well as the method used for searching the subtree.

2.5 Wireless ATM Projects

There are a couple of research projects in both academia and industry focusing on prototyping a WATM system or proposing a modified ATM protocol for WATM networks.

BAHAMA

Lucent Technologies Inc., formerly AT&T Bell Laboratories, has developed a WATM LAN system called BAHAMA (Broadband Adaptive Homing ATM Architecture) [64, 65]. The basic characteristic of the network is its *ad hoc* nature. In their proposed LAN, network nodes called Portable Base Stations (PBS) provide microcell coverage. The PBSs can be distributed in an arbitrary topology to form the backbone network of a LAN.

Handoff issues are handled in the BAHAMA project by emulating connectionless networks using ATM technology. It is similar to the concept of Mobile IP except that mobility management and handoffs are handled at the ATM layer. To emulate the connectionless services of IP, VPIs are assigned as addresses of network nodes (switches). All cells with a given VPI are routed to the same destination switch. Each connection is associated with two *Home Stations* (*Source Home Station* and Destination Home Station) in the wired network. A *Home Station* is responsible for routing the packets of a connection to their destination. When a mobile user moves from one location to another during the course of a connection, the mobile user must register with a new *Home Station*. The new *Home Station* will contact the previous *Home Station*, which is responsible for rerouting packets to the current *Home Station* and/or to the remote *Home Station*. Regarding the routing and topology management, BAHAMA employs a protocol called VTRP (*Virtual Trees Routing Protocol*) [22] to dynamically update VPI routes with changing traffic conditions and network topology. This is based on the pre-setup *destination rooted virtual trees* via ATM VPI. Advantages of this Homing Algorithm include simple control and preservation of FIFO cell sequence within a VC and between VCs with common end-points.

Columbia University

In [66], Acampora and Naghshineh propose a virtual connection tree scheme to handle the handoff problem in a WATM environment. A virtual connection tree is a collection of base stations and wired network switching nodes and links. The root of the tree is a fixed switching node of the wired network and the leaves of the tree are mobile access points or base stations. For each mobile connection, the connection tree provides a set of virtual connections (in each direction), each providing a path from the root to one leaf. At the time a mobile connection is admitted to a connection tree, the call setup procedure is executed in two steps. First, the fixed portion of the virtual connection is established between the root of the tree and the appropriate fixed point of the wired network. Second, within the connection tree, two sets of connections numbers are assigned to that mobile connection. One member of each set is used to define a path from the root to one of the leaves. The routing tables of the switches within the connection tree are appropriately updated to include the new connection numbers. The advantage of this scheme is that when a mobile user that is already admitted to a virtual connection tree wishes to handoff to another base station in the same virtual connection tree, the user simply begins to transmit packets with the connection numbers assigned for use between itself and the new base station. Using the pre-established path between the new base station at the root of the tree, the mobile's packets will flow to the root, and cross the fixed portion of the network, to their ultimate destination.

NCNR

In [25], Akyol and Cox propose a handoff scheme called *Nearest Common Node Rerouting (NCNR)*. They assume that the wireless network service area is partitioned into "zones" [67]. Each zone consists of a set of radio ports, the radio port controllers, and the fixed network infrastructure that interconnects these radio ports and provides fixed network access. The zone is managed by the "Zone Manager" process, which is responsible for database and connection management functions in the zone. Based on the zone concept, the handoff procedures can be divided into intra-zone handoff and inter-zone handoff. The intra-zone handoff is only performed in the WATM network interface equipment within the zone. This does not involve ATM network level rerouting. The inter-zone handoff occurs when the radio ports involved in the handoff belong to different zones. The NCNR scheme attempts to perform the rerouting for an inter-zone handoff at the closest ATM network node that is common to both zones involved in the handoff transaction. This scheme minimizes the resources required for re-routing and conserves network bandwidth by eliminating unnecessary connections.

RATM

RATM (Radio ATM) is the experimental WATM LAN project at Cambridge-Olivetti Research Laboratories in United Kingdom [68]. In this project, a WATM handoff protocol was developed and evaluated using Fairisle ATM switches [23]. These investigations proposed a hybrid handoff protocol based on an incremental re-establishment scheme and the concept of crossover switch discovery. The hybrid handoff protocol is designed to exploit the advantage of radio hint, with consideration given to circuit reuse efficiency, QoS, scalability, traffic continuity, traffic disruption symmetry, call loss and sequencing. This scheme requires a cross switch discovery algorithm to be performed every time a handoff session occurs. The cross switch discovery process will introduce time delay overhead to the handoff process.

SWAN

The SWAN (Seamless Wireless ATM Network) project [69], in Bell Laboratories is an experimental indoor wireless network that enables researchers to investigate the combination of wireless access and multimedia network computing, in an indoor environment. It is based on room-sized pico-cells and mobile multimedia end-points. SWAN enables users with multimedia devices such as PDAs, laptops, and portable multimedia terminals, to seamlessly roam while accessing multimedia data resident in a backbone wired network. The network model of SWAN consists of base stations connected by a wired ATM backbone network, and WATM last-hops to the mobile hosts. SWAN is one of the first systems to realize the concept of a wireless and mobile ATM network. Each mobile host as well as each base station is embedded with custom designed ATM adapter cards called FAWN (Flexible Adapter for Wireless Networking). FAWN uses off-the-shelf 2.4 GHz ISM band radios. In the design of signaling software, they assume the existence of separate signaling protocols for the wired and the wireless network respectively, with base stations acting as gateways between the two logically distinct networks. The signaling protocol in the wired network is used to establish, teardown and rebuild connections. The signaling protocol in the wireless network is used to extend routes, remove route loops and trigger route rebuilds, in addition to the establishment, teardown and rebuilding of connections. By using these protocols, SWAN allows the routing path to be extended within the same domain once a mobile host moves across a cell boundary. When a domain boundary is crossed, a rebuild of the routing path will be performed since building an extension will be very costly in terms of network resource consumption.

WATMnet

Raychauhuri and *et. al.* in NEC have been developing a prototype WATM system, called *WATMnet* [70]. The focus here is on providing a PCS data link sublayer and PCS MAC sublayer under the ATM network layer to make the wireless transmission media transparent to the ATM network layer. At physical layer, WATMnet employs the TDMA radio signal scheme. Based on TDMA, this prototype uses a Media Access Control (MAC), called Multiservice Dynamic Reservation (MDR), to provide reasonable quantitative QoS levels of B-ISDN type services. In the design of MDR, a TDMA frame is subdivided into N_r request slots and N_t message slots. Each message slot provides for transmission of a packet or ATM-like *cell* with data payload of 48 bytes (or a submultiple $48/n$ bytes, where appropriate) together with PCN protocol headers. Request slots are comparatively short and are used for initial access in slotted ALOHA (contention) mode. With this air interface, CBR service can be supported by assigning a number of fixed slots in each TDMA frame. By assigning one or more 48 byte slots in the TDMA interval following the last allocated CBR slot in the frame, it is able to support VBR traffic. Long data messages (bursts) which cannot be fitted in one frame may be segmented and rate controlled for transmission in multiple frames [29]. To deal with the high error rate problem on wireless links, a PCS packet is equipped with a 2 byte CRC field for the 48 byte payload.

The experimental system hardware is composed of laptop computers with WATMnet interface cards, multiple VME/i960 processor-based WATMnet base stations, and a mobility-enhanced local-area (2.4 Gbps) ATM switch. The prototype wireless network interface cards operate at peak bit rates up to 8 Mbps, using low-power 2.4GHz modems. Experiments with the WATMnet prototype have been conducted to validate major protocol and software aspects, including DLC, wireless control, and mobility signaling for handoff. Mobile multimedia/video applications requiring moderate bit rates (~0.5–1 Mbps) in ABR mode have been successfully demonstrated [71]. However, they only focus on system design issues involving a single ATM switch. It is not clear how they handle the user mobility problem and reroute handoff VCs over a multiple-hop environment.

2.5.1 Related Work on Handoff Management

Handoff algorithms in the literature can be classified as: end-to-end connection re-establishment, incremental connection re-establishment, multicast-based re-establishment, anchor rerouting, and virtual connection tree rerouting [22, 23, 24, 67, 72].

End-to-end connection re-establishment: When a mobile host moves to a new location, a new routing path must be computed and a new end-to-end connection will be set up [72]. Then the original connection is torn down. Computing a routing path is a time-consuming process since setting up a connection consumes a tremendous amount of network resources. The problem is even worse when the handoff frequency is high. Therefore, this scheme is not practical for multimedia applications of future WATM networks.

Incremental connection re-establishment: Unlike the end-to-end connection re-establishment scheme, the incremental re-establishment scenario reuses as much of an existing connection as possible, creating only the portion between the cross node and the new BS. The corresponding portion of the original connection is then torn down. *NCNR*, *RATM*, and *SWAN* all employ the incremental connection re-establishment scheme to handle connection reroutings. *NCNR* and *RATM* both perform crossover node discovery algorithms to determine where to establish the new partial route. *NCNR* invokes the discovery function on a per switch basis by flooding messages on the networks. *RATM* proposes several discovery algorithms and only invokes them on a per cluster basis. In terms of preserving cell sequence, both algorithms do not provide a schematic way to keep cells in sequence. Also, *RATM* does not address the issue of misrouted cells, and *NCNR* simply claims that those cells can be forwarded to the new BS without providing any details. The *SWAN* project extends the routing path of a handoff connection provided it occurs within the same domain. Although this scheme can preserve cell sequence during a connection handoff in a domain, it may violate the QoS of a handoff connection by simply extending its route. Another problem with this scheme is when a domain boundary is crossed, it requires a new routing path to be rebuilt. Rebuilding a new routing path will cause tremendous delay and consume excessive network resources.

Multicast-based re-establishment: To smooth the handoff transition, multicasting is applied to establish duplicate data streams in multiple routing paths. Thus, when a mobile moves from one BS to another BS, the data stream is already there. This will minimize the disruption gap of data streams. In addition to the incremental scheme, *NCNR* and *RATM* provide another scheme by employing multicasting mechanisms to deal with connection handoffs. They construct routing path to and from the common node to the new BS and the current BS before the handoff process stabilizes. Downlink traffic are multicasted to both BSs. Multicasting will cause data duplication at the mobile host and the crossover switch. Since ATM cells are not equipped with cell sequence numbers, handling duplicate data streams will require the cells to be reassembled into their AAL frames. It is not acceptable to perform this function on ATM switches as it would increase system complexity and transmission delay. Hence, duplicated data also means that duplicated buffer space is needed.

Anchor rerouting: Similar to the spirit of mobile IP, an anchor rerouting scheme always associates a mobile with an anchor node. When a connection is established between a mobile and a remote server, it must include the mobile's anchor node. All data sent to the mobile from the source must be delivered to the anchor node. The agent on the anchor node forwards the data to the mobile. The *BAHAMA* project employs this scheme to cope with the highly dynamic topology of ad-hoc networks. This scheme is known as having the tri-angle routing problem. That is, the derived routing path is usually much longer than the direct routing path between the source and destination. Therefore, this scheme may be feasible for wireless ad-hoc networks, it is not suitable for WATM networks, as longer routes will result in more network resource consumption and QoS violations. Another problem of this scheme is that it does not scale well on large networks due to VPIs exhaustion upon constructing the rooted virtual tree. Hence it wastes wired network bandwidth while broadcasting location queries. However, cell sequencing can be preserved since the anchor agent will take care of this before forwarding cells to their destination.

Virtual connection tree routing: The virtual connection tree concept is proposed by Acampora at Columbia University. As described in the previous sections, a virtual connection tree must be established when a connection is accepted in a domain. By doing this, multiple virtual circuits must be established to support a single connection. This tends to exhaust the VCIs in the ATM networks. There is no cell sequence guarantee in this scheme. Also, this scheme does not handle misrouted cells.

2.6 Summary

In this chapter, we presented background information, such as wireless technologies, ATM technologies, and PNNI routing, for the study of WATM networks. We also reviewed related works on handling user mobility location in the literature.

TDMA and CDMA are the two candidate technologies for transmitting ATM traffic over wireless links. The major benefit of using TDMA is that the technology is simple and mature. CDMA, on the other hand, is relatively new. The near-far problem of CDMA systems requires complicated power control mechanisms to be implemented. But CDMA systems have several benefits over TDMA systems, including supporting multiple user on the same wideband channel, soft handoff, and dynamically adjusting wireless communication bandwidth.

The VPI/VCI concept in ATM technologies allows ATM networks to inherit benefits from both circuit switching and packet switching schemes without including their problems. End-to-end virtual circuits allow data to be delivered to their destinations quickly along fixed routing paths. Statistically multiplexing multiple virtual circuits in a virtual path avoids low utilization of dedicated circuits due to inactive communications. The use of VP also enables us to construct virtual networks over a physical ATM network.

Location management and handoff management are the two major tasks in handling user mobility in wireless networks, especially in the future WATM network. Although IS-41 location management is widely used in today's PCN, it is not efficient in tracking user location and exchanging signaling messages. Many improvements of IS-41 location management have been proposed, including per-user location caching, user profile replication, pointer forwarding, local anchoring, tree-base distributed database, and partitioning schemes. Per-user location caching, user profile replication, pointer forwarding, and local anchoring are based on the IS-41 location management. Each scheme aims to reduce the signaling overhead by minimizing the number of signaling message exchanges, decreasing the travel distance of messages, or reducing the frequency of database lookup. Tree-base distributed database approaches and partitioning schemes provide new location management mechanisms instead of improving the original IS-41 model. Tree-base distributed database approaches need special arrangement of database so as to reduce the number of database accesses and updates. The partitioning scheme requires the knowledge of the mobility pattern to partition registration location efficiently and to reduce the frequency of location updates.

Research on providing handoff management in WATM networks has been very active in last few years. We surveyed current available WATM handoff management schemes and discussed their tradeoffs. These handoff schemes can be classified as: end-to-end connection handoff, incremental connection handoff, multicast-based handoff, anchor rerouting, and virtual connection tree routing. An end-to-end connection handoff is not practical for WATM networks since it incurs prohibitive connection construction overhead. Incremental connection handoff schemes can reuse most parts of an original routing path in order to minimize the overhead of constructing a new routing path. Nevertheless, the efficiency of this scheme depends on the network architecture as well as the rerouting mechanism. Although a multicasting approach may support smooth connection handoffs by duplicating data streams to multiple BSs, it increases the system complexity and network bandwidth consumption. Anchor rerouting schemes can preserve cell sequence but they tend to violate the QoS of a handoff connection. Finally, the virtual connection tree scenario reduces the delay of a connection re-establishment at the cost of exhausting VCIs. The drawback of this scheme is that it can not preserve the cell sequence of a handoff connection.

Chapter 3

Framework for WATM Location Management

In this chapter, we present the framework for the WATM location management. The framework includes a WATM network architecture and a WATM location management scheme. This location management scheme is designed especially for WATM networks and is tightly coupled with the proposed WATM network architecture, to support efficient location management functionalities.

3.1 WATM Network Architecture

WATM network architectures are essential for integrating wireless networks with wireline ATM networks. A well designed WATM network architecture should have the following features:

- cooperate with current deployed ATM networks,
- minimize the modifications to be made on the ATM standards,
- perform location management functions efficiently,
- handle connection handoffs smoothly,
- reduce the PCS migrating cost.

As shown in Figure 3.1, we propose a WATM network architecture, called **Virtual Path WATM Architecture (VPWA)** [73, 74], to satisfy features listed above. The WATM network architecture consists of a core ATM/B-ISDN backbone network and peripheral network structures for wireless communications. When constructing a WATM network, the separation of the core ATM backbone network and the wireless network structure allows mobility support functions to be implemented on the peripheral network without major modification on the deployed ATM network. This will avoid the tremendous cost of modifying or replacing deployed ATM switches. Therefore, a smooth integration of an ATM backbone network and wireless PCS networks is possible.

The VPWA makes use of the virtual path concept of ATM technologies to construct a virtual network topology on top of a fixed ATM infrastructure. The main advantage of a hierarchical network architecture is that routing is simple and straight forward. Another advantage of adopting a hierarchical network infrastructure is that it is easy to implement the topology aggregation for a hierarchical routing algorithm, which reduces complexity in topology advertisement, and provides fast connection rerouting [75]. On the other hand, because of the use of the VP technology, the virtual network structure can be easily re-arranged as desired. Therefore, the problem of a rigid physical network infrastructure in a hierarchical network can be avoided.

The VPWA scheme applies the clustering technique proposed in [24]. A cluster includes a group of neighboring BSs and ATM switches that connect these BSs. Each cluster selects an ATM switch as its cluster coordinator, termed a Mobile Cluster Switch (MCS). A MCS is an ATM switch with desired mobility functions built into it. Generally, a cluster may simply consist of an ATM switch and the BSs attached to it. But the present scheme reserves the flexibility to expand the coverage area of a cluster by including multiple ATM switches. In this case, the select MCS may not have direct links to all BSs in a cluster. Nevertheless, it would still appear as if all BSs have direct connections to their MCS because of the use of VP. Every BS in a cluster has VPs connected to and from the MCS. A BS also has VPs to linked each of its neighboring BS. The VP between neighboring BSs supports our proposed connection handoff mechanism, which is described in Chapter 4.

Together, several MCSs constitute a domain, with a Mobile Gateway Switch (MGS) in each domain. Again, like all BSs in a cluster, every MCS in a domain has VPs connected to and from its MGS. A MGS functions as the gateway to the ATM/B-ISDN backbone network for all MCSs in that domain. Note that the MCS at the boundary between two domains may belong to multiple domains with VPs to those MGSs. We call such a MCS a *border MCS*. This special arrangement is used

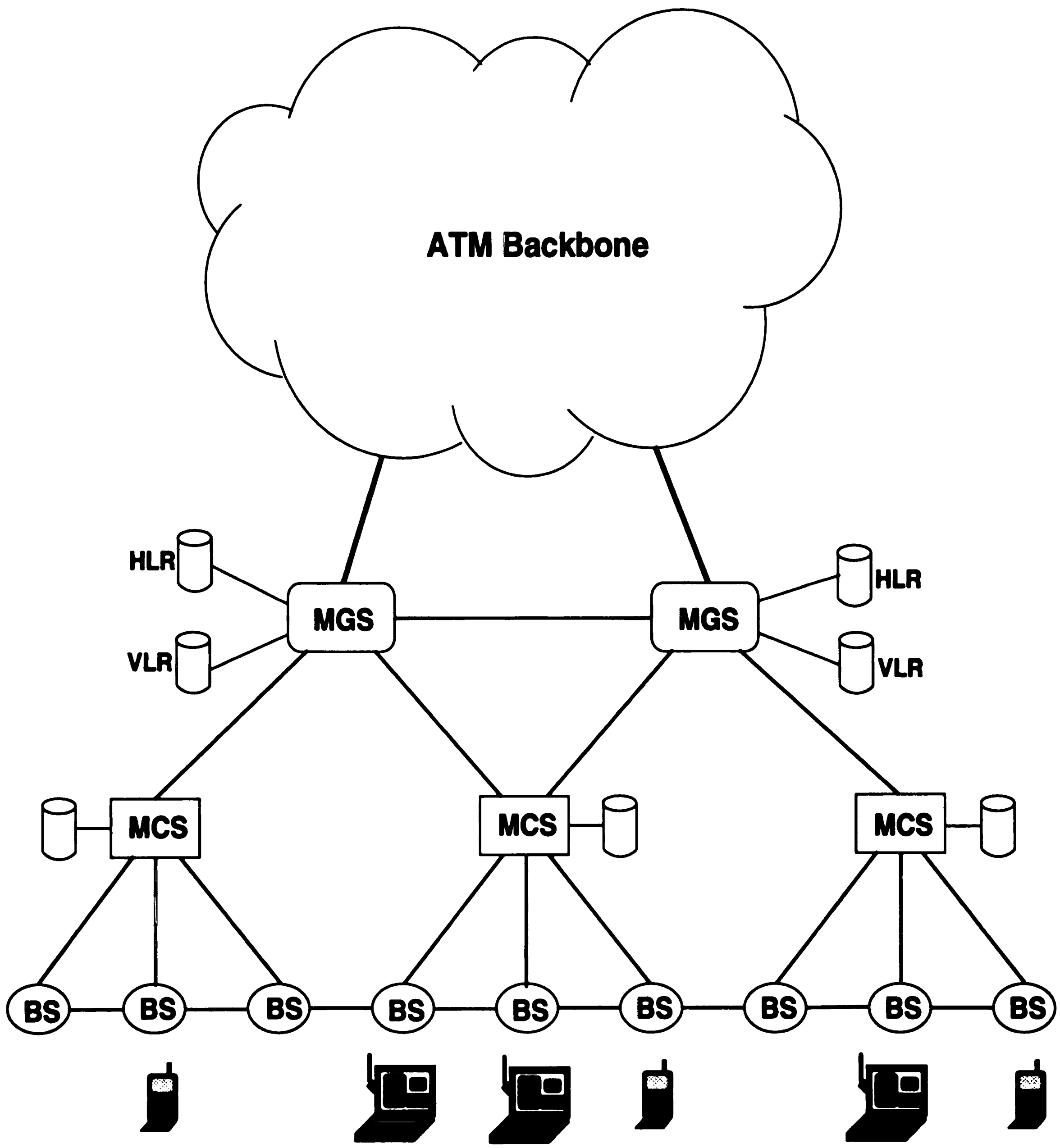


Figure 3.1: Virtual path wireless ATM architecture

to relieve the thrashing problem in location management and handoff management caused by a MH moving forward and backward between neighboring domains. Further details about thrashing of location management is presented in the next section. Detailed procedures of handling thrashing in handoff management are presented in Chapter 4.

3.2 WATM Location Management

Consistent with the proposed WATM network architecture, we construct a new location management scheme, called VPWA location management, for the future WATM network. Although the surveyed location management schemes in Chapter 2 are not suitable for the WATM network since they are designed for the current PCS network, some of the techniques can be applied in the design of the location management scheme for a WATM network. Our location management scheme is a hybrid scheme which adopts the HLR-VLR scheme in IS-41 [21], the distributed database approach in [61], the pointer forwarding scheme in [59], and local anchoring in [60].

We set the domain in the VPWA equal to a PNNI peer group at PNNI level P, with the MGS as the group leader. Every MH is associated with a home domain. It maintains a Permanent ATM NSAP (PAN) address (assigned by its home domain) as its identification in the VPWA network. A *border MCS* has multiple ATM addresses, each belonging to a domain in which the MCS is involved. As in the IS-41 model, each MGS switch is equipped with a HLR database and a VLR database. The

VPWA location management scheme improves upon the IS-41 model by using the forwarding pointer technique to reduce the number of queries to the HLR. However, unlike forwarding pointer scheme, the VPWA uses one-hop pointers. No pointer chains are allowed since a pointer chain results in a long call setup delay due to the cranking back effect on the ATM network. Also the distributed database approach is used in each domain. In additional to the HLR and VLR databases on a MGS, a MCS switch is equipped with a local database, which keeps track of all mobile hosts within its cluster. The distributed database is impractical in a WATM network because it conflicts the PNNI routing mechanism by constructing a hierarchical routing tree on the entire network.

Before further discussing our WATM location management scheme, we define a few abbreviations to simplify its description. We refer to a *home MGS* as the MGS of a MH's home domain and a *visiting MGS* as the MGS of the domain visited by a MH. A *home MCS* means the MCS of the cluster which is currently serving the MH when the MH is in its home domain. A *visiting MCS* is the MCS of the cluster which is currently serving the MH when the MH is in a visiting domain. A new base station (NBS) is the base station that will serve a MH after a movement of the MH. A current base station (CBS) is the base station which serves a MH before it moves.

3.2.1 Location Tracking

When a MH powers up **inside its home domain**, it initiates a registration process by sending a registration message to its serving BS. The registration message must include the PAN address of the MH. A user identification (UID) number and/or a calling number which is used in the current PCS network may be included in the signal message for migrating PCS services to WATM networks. The serving BS forwards the registration to its MCS. Upon receiving the registration message, the MCS simply updates its local database to record the MH's current location. The registration signal is further forwarded to the MGS by the MCS to update the HLR database. A HLR database entry keeps the PAN address and the UID of the MH, and the ATM address of the serving MCS.

When a MH powers up **outside its home domain**, the steps shown in Figure 3.2 are performed to keep track of the location of the MH. These steps are described below:

1. A MH initiates a location registration message to the serving BS (*BS A.1.3*).

2. The registration signal is forwarded to the serving MCS (*MCS A.1*) via the BS. Upon receiving the registration message, the MCS examines the message and determines that this MH is a visitor. A database entry is created in the local database to record the PAN address and the UID of the MH. The MCS then forwards this message to its MGS (*MGS A*).

3. When *MGS A* receives the message, it generates a temporary ATM NSAP (TAN) address for the visiting MH. The TAN address and the PAN address and the UID of the MH forms the tuple (TAN, PAN, UID), which is recorded in the VLR of *MGS A*. The MGS sends an acknowledgment with the TAN

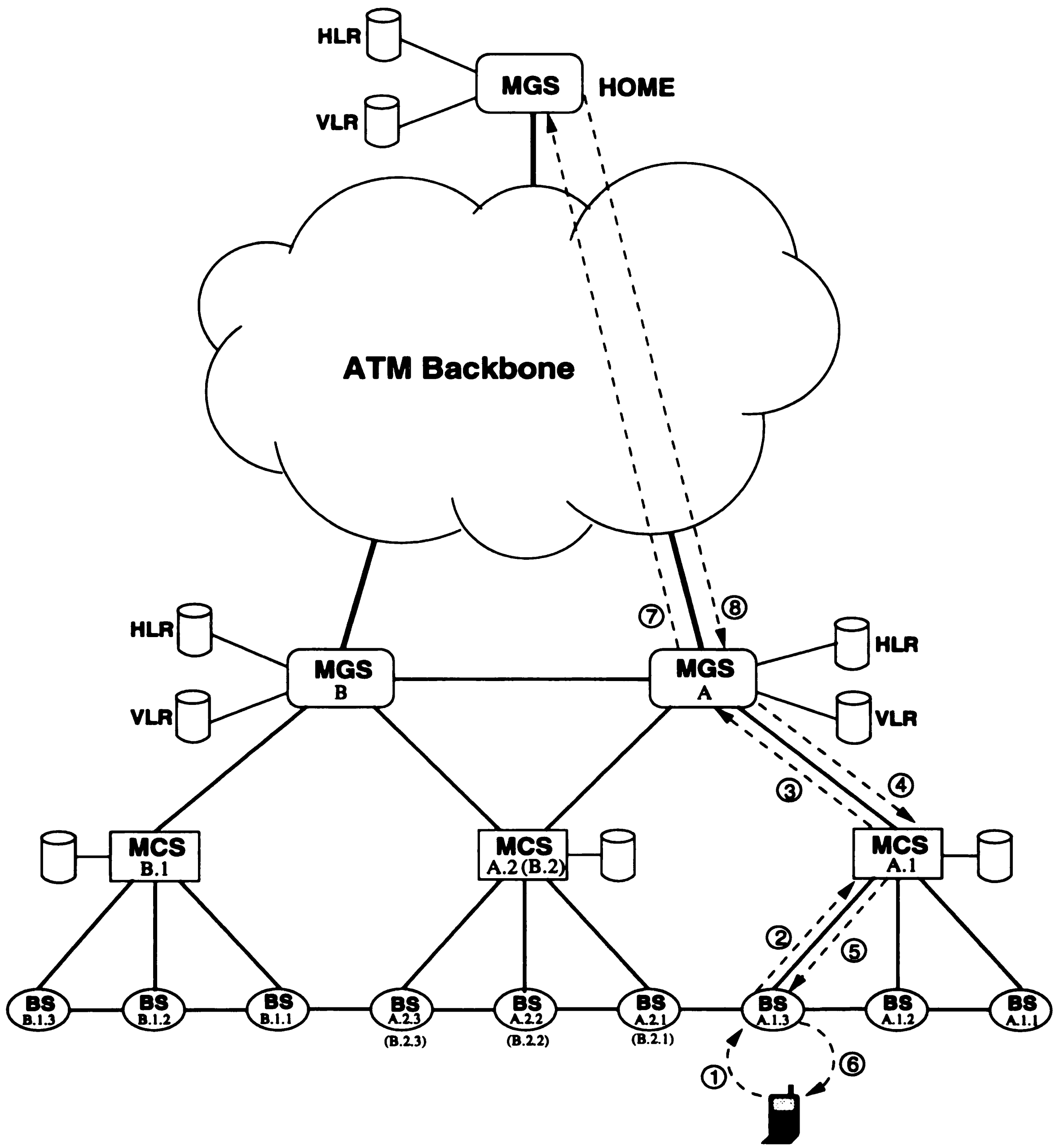


Figure 3.2: Location registration in a Wireless ATM network

address back to *MCS A.1*.

4. When *MCS A.1* receives the message, its records the tuple (TAN, PAN, UID) in its local database. The acknowledgment and the TAN address is further forwarded to *BS A.1.3*.

5. *BS A.1.3* sends this acknowledgment to the MH.

6. When the MH receives the TAN address, it keeps the address for its temporary identity in the domain.

7. *MGS A* also sends an update message with its ATM NSAP address to the *home MGS* of the MH. When the *home MGS* receives the location update signal, it changes the forwarding pointer of the MH to *MGS A* in the HLR. It also sends the user profile of the MH with an acknowledgment back to *MGS A*.

8. Immediately after *MGS A* receives the user profile from the *home MGS*, its stores the user profile in its VLR.

When a mobile moves to another BS in the same cluster, only the local database of the *visiting MCS* is updated to point to the serving BS. No update is needed at the *visiting MGS*. If the MH moves out of a cluster, the location update procedure illustrated in Figure 3.3 is executed to keep track of the location of the MH. Steps in the procedure are listed below:

1. A MH moves to a new cluster and sends a registration update message to the new BS (*BS A.2.1*). *BS A.2.1* forwards the message to its MCS (*MCS A.2*).

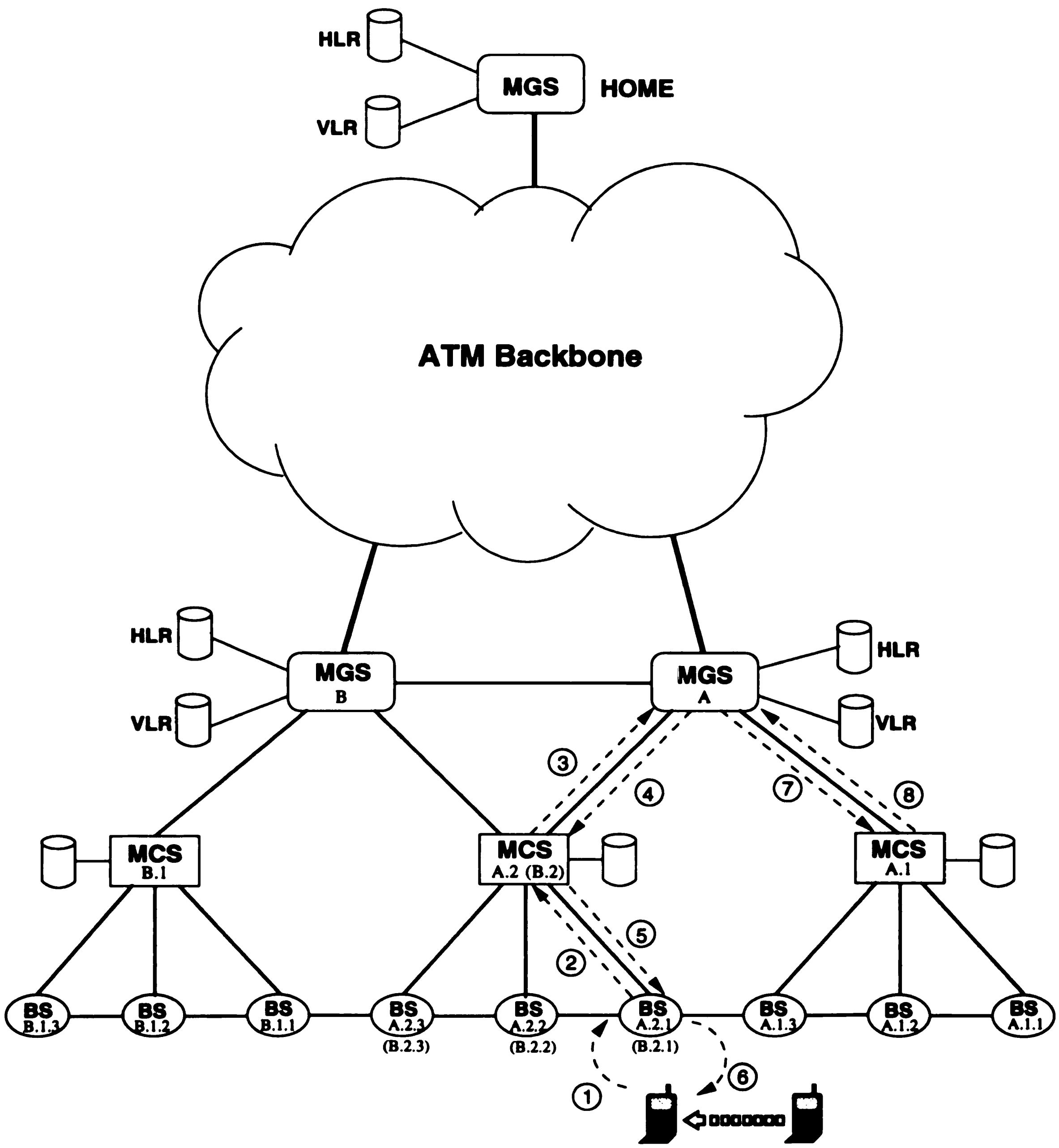


Figure 3.3: Inter-cluster location tracking in a Wireless ATM network

2. Upon receiving this signal, *MCS A.2* stores the PAN address and the UID of the MH in its local database. The registration update message is forwarded to *MGS A* by *MCS A.2*.

3. When *MGS A* receives the signal, it determines that the MH moves to a new cluster by checking its VLR. The pointer in the VLR is updated by pointing to *MCS A.2*. This MGS also sends an acknowledgment back to *MCS A.2*. It also sends a cancellation signal to the previous MCS (*MCS A.1*).

4. When *MCS A.2* receives the acknowledgment from its MGS, it forwards the message to *BS A.2.1*.

5. *BS A.2.1* further forwards the acknowledgment to the MH upon receiving the signal.

6. When the MH receives the acknowledgment, it knows that it has registered itself in the new cluster.

7. When *MCS A.1* receives the cancellation signal from *MGS A*, it deletes the entry from its local database.

8. *MCS A.1* sends an acknowledgment back to *MGS A* to confirm that the database entry for the MCS is removed.

Note that in Figure 3.3, the *MCS A.2 (B.2)* is a *border MCS*. When a MH moves from BS A.1.3 to *BS A.2.1*, the inter-domain location registration procedure is not

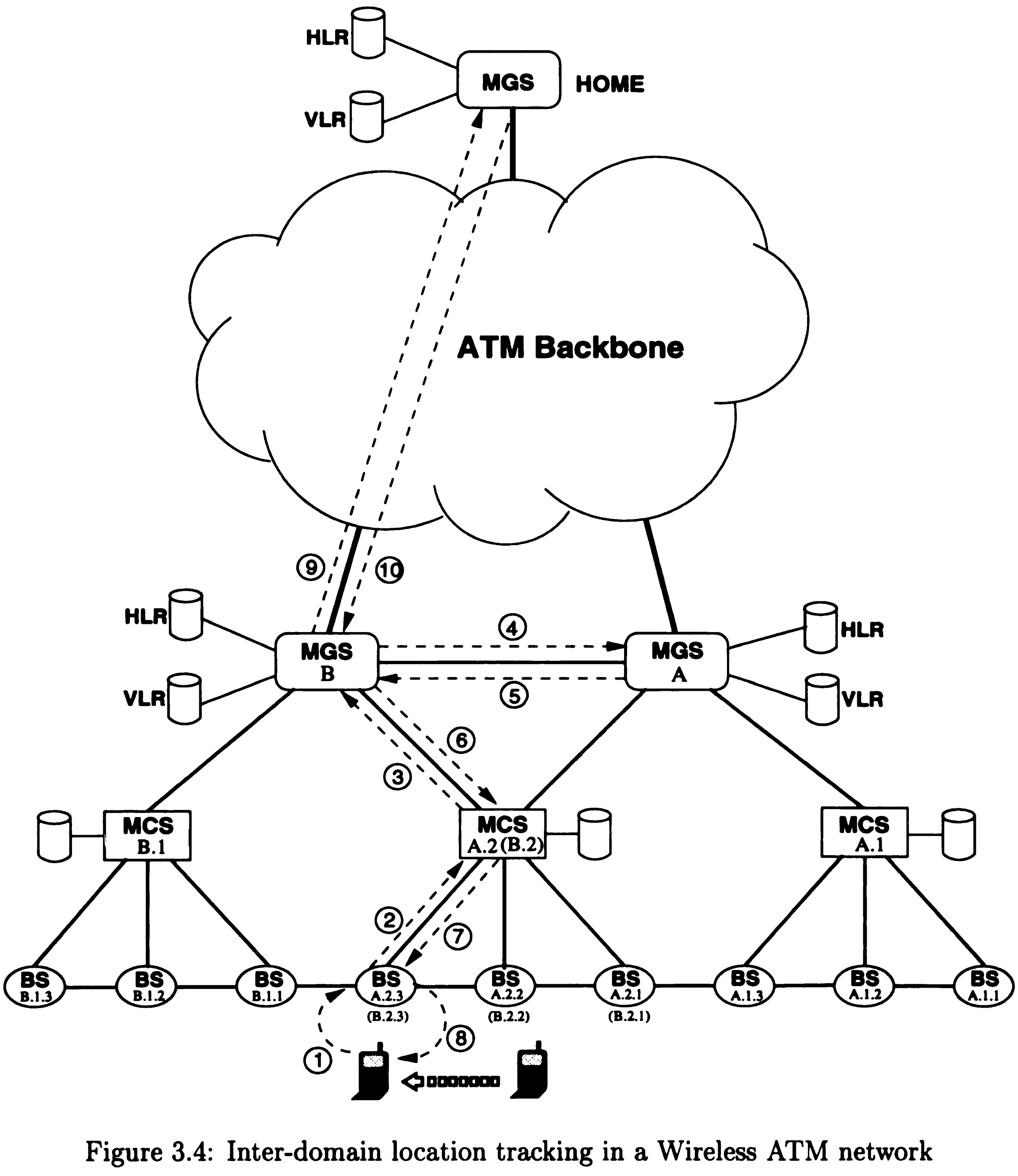


Figure 3.4: Inter-domain location tracking in a Wireless ATM network

invoked right away. Instead, an inter-cluster location update is performed. An inter-domain location update procedure is executed only when the MH moves deeper into the cluster of the *border MCS*. For example, in Figure 3.4, if the MH moves from *BS A.2.2* to *BS A.2.3*, the inter-domain location tracking is invoked. This scheme uses the coverage area of the *border MCS* as the buffering zone for reducing location updates between neighboring domains. Therefore, the signaling and processing overheads incurred by location updates are substantially reduced. Detailed message flow of Figure 3.4 are explained in the following steps:

1. When a MH moves from *BS A.2.2* to *BS A.2.3*, the MH sends a registration update request to *BS A.2.3*.

2. *BS A.2.3* forwards the request to *MCS A.2*. Upon receiving the request, the MCS decides to initiate an inter-domain registration update procedure because the MH is moving deeply into the domain B. It sends a location registration signal to *MGS B* along with the tuple (TAN, PAN, UID) of the MH in domain A.

3. Immediately after receiving the registration signal, *MGS B* assigns a TAN address to replace the old TAN address in the tuple and records the new tuple (TAN, PAN, UID) of the MH in its VLR. Based on the old TAN address, *MGS B* determines the previous serving domain (domain A) and sends *MGS A* a signal to request for the user profile of the MH, set a forwarding pointer, and cancel the location registration of the MH.

4. Upon receiving the signal from *MGS B*, *MGS A* removes the database entry of the MH from its VLR, sets a forwarding pointer to *MGS B*, and sends the user profile of the MH along with an acknowledgment to back to *MGS B*.

5. *MGS B* records the user profile of the MH in its VLR upon receiving the message. It then sends an acknowledgment along with the new TAN address to *MCS B.2*.

6. Immediately after *MCS B.2* receives the new TAN address, it replaces the old TAN address by the new TAN address. It also forwards the message to *BS B.2.3*.

7. Upon receiving the message, *BS B.2.3* delivers the message to the MH.

8. When the MH receives the message, it updates its TAN address, and the MH now belongs to domain B.

9. When *MGS B* receives the acknowledgment from *MGS A*, it also sends a pointer update message to the *home MGS* of the MH.

10. When the *home MGS* receives the update request, its updates the forwarding pointer in its HLR to point to *MGS B*. So the *home MGS* can correctly answer queries for the MH.

Note that to reduce the signaling and the database lookup overhead, a mechanism similar to the remote pointer strategy in [76] can be used. The basic idea of the strategy is that forwarding pointers can be set at selected domains, based on the call arrival rate and the mobility. Therefore, calls generated in a domain with a forwarding pointer for the serving MGS of the called MH can be delivered without consulting the HLR of the *home MGS*. This not only eliminates the database lookup at the HLR but also reduces the number of signal exchanges between the initiating MGS and the *home MGS* and thus the delay of a call setup. However, by using this scheme, all the forwarding pointers must be updated when each inter-domain movement of a MH is occurred. That is, when *MGS B* in Figure 3.4 updates the forwarding pointer at the *home MGS* in message 7, it must also update all forwarding pointers at the selected domains.

The major advantage of this proposed scheme is that, like most improvement schemes of IS-41, it reduces the traveling distance of signals and minimizes the number of signals issued. Hence, this location tracking scheme is designed for the WATM network instead of the PCN network. By using the distributed database approach within a domain, signal exchanges for an intra-cluster movement of a MH are confined within a cluster. Signals only travel up to the MGS of the domain if a MH migrates within a domain but between different clusters.

Further, the MGS of a domain functions as a dynamic anchor for handling queries of MHs in the domain. Only the HLR keeps a pointer to the serving MGS. This allows this scheme to inherit the benefits from a local anchoring scheme, which eliminates the need to report location changes to the HLR. When a MH moves into a new domain, the anchor point is changed to the new MGS. An update signal is sent to update the pointer on the HLR to the new MGS. Also, the new MGS sets a forwarding pointer on the old MGS to forward location queries of the MH. Last, the use of a *border MCS* can relieve thrashing problems among neighboring domains. Therefore, unnecessary signaling overhead, caused by a MH moving back and forth between neighboring domains, is eliminated.

3.2.2 Call Setup

To setup a call in a VPWA network, the location of the called MH must be determined first before an end-to-end connection can be made. By using the location tracking schemes described above, the current location of a MH is available by querying databases. The call setup process is different from the standard ATM connection setup procedure, which has no location determination phase before the connection setup phase. According to the VPWA WATM architecture, a call setup process can be either an intra-domain or an inter-domain call setup.

Intra-Domain Call Setup

Setting up a call within the same domain is straightforward via the proposed VPWA architecture. When a call is generated for a MH within the same domain, the calling MH first sends a **SETUP** message to its serving BS. The message is forwarded to the serving MCS. The MCS determines whether the called MH is located in its cluster by examining the local database. If the target MH is in the same cluster, the MCS simply forwards the call request to the BS which is serving the called MH. However, if the user profile of the MH is not in the local database, it must be downloaded from the VLR of the serving MGS. Unlike the paging process in the current PCS network to determine the serving BS, the MCS has up-to-date location information of all MHs in its local database. This avoids the delay of the paging process when setting up a call. If the called MH does not currently reside in the cluster, the MCS forwards the call request signal to its MGS. The MGS checks both the HLR and VLR databases and determines the location of the called MH. The call request is forwarded along the hierarchy down to the target BS and then to the called MH, and the call setup within the same domain is thus completed.

Inter-Domain Call Setup

An inter-domain call setup is more complicated than an intra-domain call setup. Every MGS functions as a local anchor point for setting up a call. That is, an inter-domain call setup is delivered to the target MGS first. From there, the call setup request is delivered to the called MH by database lookups. Figure 3.5 shows an inter-domain call setup procedure. The procedure is described in the following steps:

- Step 1: When a MH at *BS B.1.2* initiates a call to a MH at *BS A.1.1*, it first sends a **SETUP** signal to *BS B.1.2*.
- Step 2: Upon receiving the call request, *BS B.1.2* forwards the request to *MCS B.1*.

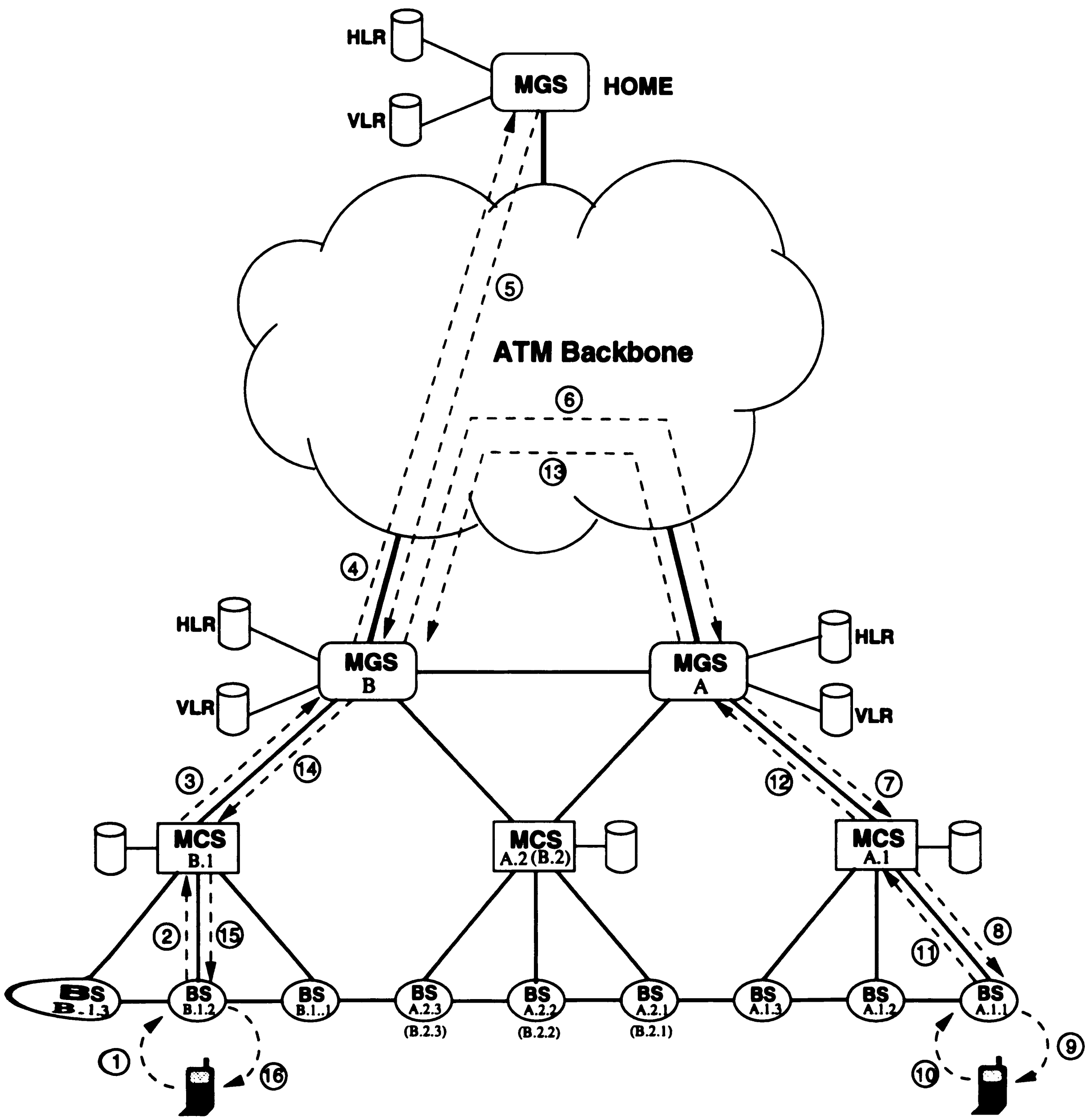


Figure 3.5: Inter-domain call setup procedure in a Wireless ATM network

- Step 3: *MCS B.1* consults its local database and finds that the called MH is not in its cluster. It then forwards the call request signal to *MGS B*. *MGS B* examines the received **SETUP** request and determines that the called MH is not in its domain. If a forwarding pointer for the called MH is available, the request is sent to the serving MGS (*MGS A*) in Step 6 without the involvement of Step 4 and 5.

- Step 4: If a forwarding pointer for the called MH is not available, *MGS B* initiates a search process to find the current location of the called MH. This is done by sending a query to the *home MGS* of the called MH.

- Step 5: When the *home MGS* receives the query, it performs a database lookup and determines that the MH is not in the home domain. The address of the *visiting MGS*, which is currently handling the called MH, is returned from the database. In Figure 3.5, *MGS A* is the current *visiting MGS* of the called MH. The address is included in the response message sent by the *home MGS* back to *MGS B*.

- Step 6: When *MGS B* receives the address of the *visiting MGS*, it forwards the call request to *MGS A* according to the standard ATM PNNI routing protocol. Note that here the destination address is changed to the address of *MGS A* and the PAN address of the called MH is moved to the destination sub-address field in the **SETUP** message.

- Steps 7–9: When *MGS A* receives the call request, it locates the called MH by checking its VLR database. Before the call request message is forwarded to the called MH, the destination address of the call request is replaced by the TAN address of the called MH. The user profile of the called MH is downloaded from *MGS A* to *MCS A.1* in Step 8 if the local database of *MCS A.1* does not have the user profile. The call request reaches the called MH at *BS A.1.1* through steps 7–9.

- Steps 10–16: When the called MH accepts the **SETUP** request, it sends a **CONNECT** back to the calling MH in the reverse direction. The delivery of the signal is performed by Steps 10–16.

The steps presented above manage call setup in general cases. Now consider what may happen if the called MH moves to another domain before a call setup arrives at the prior *visiting MGS*? Apparently, the call could be delivered to the wrong domain. However, since the new *visiting MGS* will set a forwarding pointer at the VLR of the old *visiting MGS* to forward the call request (refer to Figure 3.4). Thus, when a **SETUP** arrives at an old *visiting MGS*, the request will be efficiently forwarded to the new MGS by substituting the destination address with the address of the new *visiting MGS*. When the request is received by the new *visiting MGS*, the same call setup process within a domain is followed.

Note that when a MGS issues a location query or a location update to the home domain of a MH, it requires connection setups and releases. This may generate consid-

erable processing overhead. This problem can be relieved by applying a connectionless ATM service as in [77], which is able to provide low cost of data communications, with a service similar to the connectionless service on the current TCP/IP network.

For an intra-domain call setup, the proposed call setup scheme can quickly locate the called MH and accurately deliver the call. No query to the HLR of the called MH is necessary. For an inter-domain call, setting up the segment of the call between the calling domain and the called domain follows the standard ATM PNNI routing across the ATM backbone. This avoids major modifications on the ATM backbone network. Therefore, the proposed call setup scheme can co-exist with the current call setup of the ATM standard. This helps to migrate PCS services to a WATM network.

3.2.3 Performance Analysis

In this section, we will derive and compare the signaling cost of our VPWA location management scheme with the IS-41 scheme. We assume that a cluster in our VPWA architecture is equal to the coverage area of a mobile switching center in IS-41. Following the analysis technique in [76], several parameters are defined in Table 3.1 to be used in the analyses for the signaling cost of the VPWA and IS-41 schemes. The signaling cost can be further divided into location registration cost and call setup cost. Hence, depending on the movement of a mobile host, the signaling cost for various cases differs, as indicated in Table 3.2.

According to the parameter set in Table 3.1, the location registration cost for inter-cluster and inter-domain movements are $(1-q)\lambda_m\beta_1$ and $q\lambda_m\beta_2$, respectively.

Parameters	Definition
λ_c	Call arrival rate
λ_l	The arrival rate for calls from the same residing domain of the called MH
λ_h	The arrival rate for calls from domains without forwarding pointers
λ_m	MH mobility rate
c_{s1}	Signaling cost for sending signals between MCS and MGS
c_{s2}	Signaling cost for sending signals between HLR and MGS
c_{s3}	Signaling cost for sending signals between MGS and MGS
c_{s4}	Signaling cost for sending signals between neighboring MGSs
q	Inter-domain movement probability
v_l	Fraction of incoming calls that are from the local domain
v_r	Fraction of incoming calls that are from domains with forwarding pointers
n	Number of domains that have forwarding pointers
C_{loc}	Signaling cost for location registration in VPWA
C_{setup}	Signaling cost for call setup in VPWA
V_{cost}	Total signaling cost of the VPWA scheme
I_{cost}	Total signaling cost of the IS-41 scheme

Table 3.1: Definition of analytical parameters

The process for updating forwarding pointers is invoked only when an inter-domain movement occurs. The cost for updating forwarding pointers at remote domains is $nq\lambda_m\gamma$. Thus, the total cost for location registration is:

$$C_{loc} = (1-q)\lambda_m\beta_1 + q\lambda_m\beta_2 + nq\lambda_m\gamma \tag{3.1}$$

Similar to the location registration procedure, the signaling cost for setting up inter-cluster calls, inter-domain calls without forwarding pointers, and with forwarding pointers is $\lambda_l\alpha_1$, $\lambda_h\alpha_2$, and $\lambda_r\alpha_3$ (where $\lambda_l = \lambda_c v_l$, $\lambda_r = \lambda_c v_r$, and $\lambda_h = \lambda_c - \lambda_l - \lambda_r$),

Cost	Definition	Expression
α_1	Cost for call setup if the called MH is residing in the same domain.	$4c_{s1}$
α_2	Cost for call setup if no forwarding pointer is available.	$2(2c_{s1} + c_{s2} + c_{s3})$
α_3	Cost for call setup if a forwarding pointer is available.	$2(2c_{s1} + c_{s3})$
β_1	Cost for location registration after an inter-cluster and intra-domain movement.	$4c_{s1}$
β_2	Cost for location registration after an inter-domain movement.	$2(c_{s1} + c_{s2} + c_{s4})$
γ	Cost for updating a forwarding cost.	$2c_{s3}$

Table 3.2: Definition and expression of cost for various conditions

respectively. Therefore, the cost for call setups is:

$$C_{setup} = \lambda_l \alpha_1 + \lambda_h \alpha_2 + \lambda_r \alpha_3 \tag{3.2}$$

From equations (3.1) and (3.2), we obtain the total signaling cost for the VPWA location management scheme as indicated below:

$$V_{cost} = C_{loc} + C_{setup} \tag{3.3}$$

$$= (1-q)\lambda_m \beta_1 + q\lambda_m \beta_2 + nq\lambda_m \gamma + \lambda_l \alpha_1 + \lambda_h \alpha_2 + \lambda_r \alpha_3 \tag{3.4}$$

Following the same set of parameters in Tables 3.1 and 3.2, the total signaling cost for IS-41 location management model is:

$$I_{cost} = (\lambda_m + \lambda_c)(4c_{s1} + 4c_{s2}) \tag{3.5}$$

To compare the VPWA signaling cost with the IS-41 signaling cost, we define the relative cost ratio as V_{cost}/I_{cost}. If the relative cost is 1, both schemes have the same signaling cost. We also define the call-to-mobility ratio (CMR) as the ratio of the call arrival rate to the mobility rate. In other words, $CMR = \lambda_c/\lambda_m$. We assume that the inter-domain movement probability, q, is 0.125 and the number of domains with forwarding pointers, n, is four.

	Set1	Set2	Set3	Set4	Set5
c_{s1}	1	1	1	1	1
c_{s2}	5	10	10	10	10
c_{s3}	2	2	5	10	10
c_{s4}	2	2	2	2	5

Table 3.3: Signaling cost parameter sets

We first evaluate the signaling cost by applying the five parameter sets in Table 3.3. Because the MGS is located close to its associated MCSs, signaling messages exchanged between a MGS and one of its MCS are confined within a domain and are expected to have the smallest value. We define the value for such signals (c_{s1}) to be 1 and all other types of signals are normalized to c_{s1}. Moreover, the normalized cost of the signal c_{s2} or the signal c_{s3} is expected to be very large, since these types of signals traverse across the ATM backbone network. In general, the signaling cost between two neighboring domains should be smaller than (or sometimes equal to) the signaling cost between two random domains. The selected data sets allow us to study the effect of varying the cost parameters c_{s2} and c_{s3} on the performance of the VPWA location management scheme.

Based on the parameter sets in Table 3.3, Figures 3.6 and 3.7 show the relative

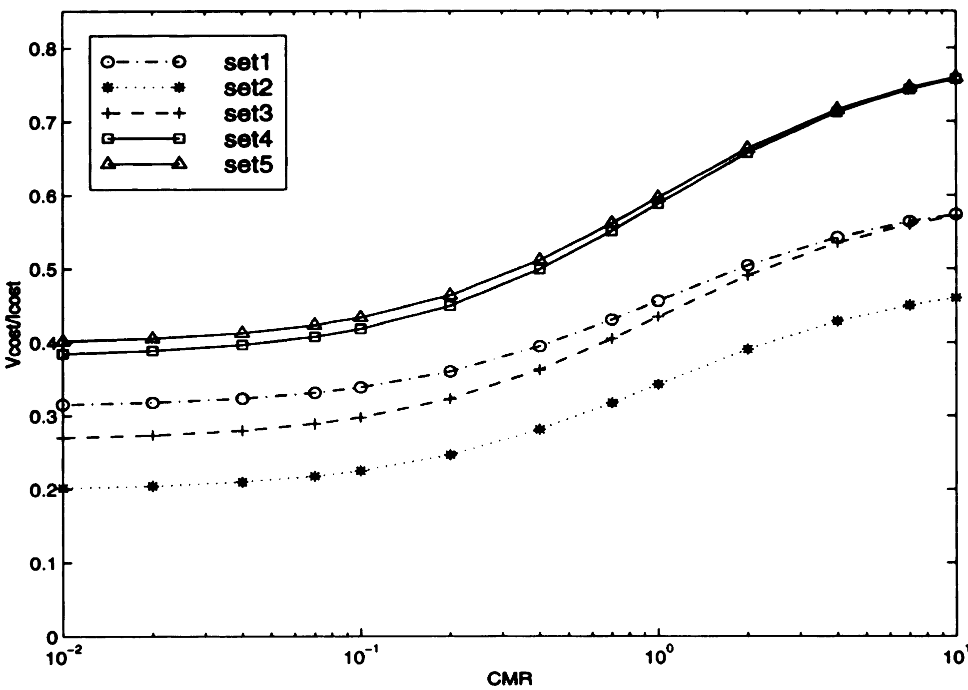


Figure 3.6: Signaling cost for $v_l = v_r = 0.15$

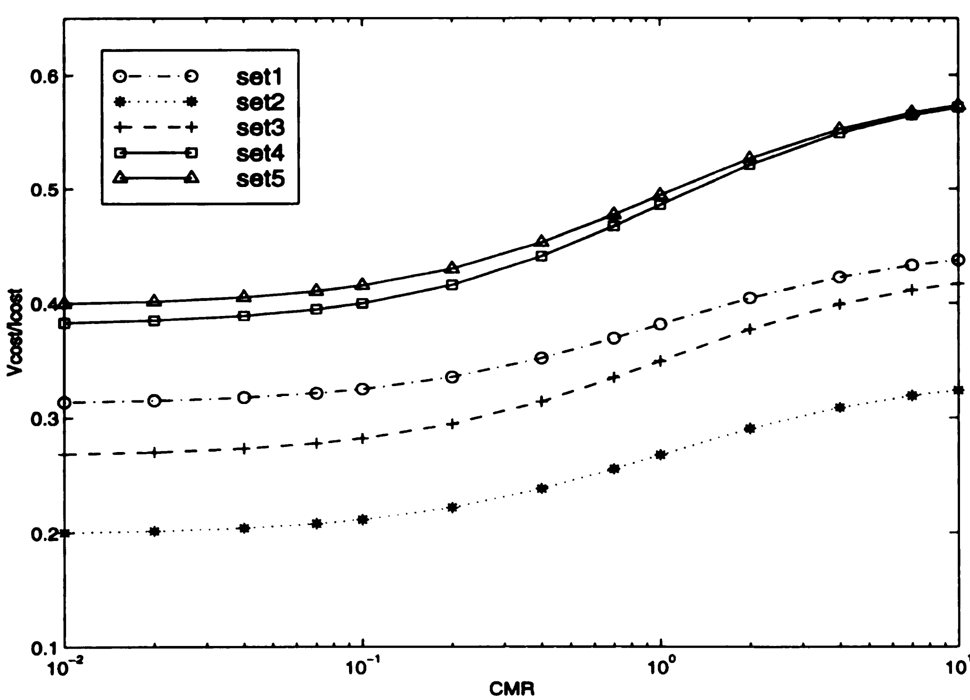


Figure 3.7: Signaling cost for $v_l = v_r = 0.30$

signaling cost for $v_l = v_r = 0.15$ and $v_l = v_r = 0.30$, respectively, when the CMR value varies in the range of 0.01–10. As it can be seen from these two figures, the relative cost increases with the CMR. When the CMR ratio is small, the mobility location registration cost dominates. The significant signaling cost reduction is due to the fact that the signals for reporting the intra-domain movements back to the HLR are eliminated. When the CMR ratio is large, the mobility rate is low relative to the call arrival rate and the cost saving from the location registration diminishes. In other words, the cost for call setups dominates and the saving in call setups become significant. This signaling cost reduction is achieved because that (1) the use of forwarding pointer eliminates the signals for querying the HLR, and (2) the intra-domain calls avoid both the HLR queries and the cross ATM backbone call setups. Thus, when both the fractions v_l and v_r increase, the signaling cost decreases as shown in Figures 3.6 and 3.7.

By comparing data set 1 with data set 3, the effect of varying the signaling cost for updating the HLR (c_{s2}) is significant when the mobility rate is high (i.e., the CMR value is small). Comparisons of data set 4 and data set 5 in both figures show that the effect of varying the signaling cost between neighboring domains (c_{s4}) become less significant when the CMR ratio is large. Clearly, this is because when the mobility rate is small, the saving from the signaling cost between two neighboring domains becomes small.

3.3 Summary

In this chapter, we have presented the WATM location management framework, which includes a WATM network architecture (VPWA) and a location management scheme. The proposed VPWA architecture consists of two parts: an ATM backbone network and a mobile-enabled peripheral network. The separated network structure allows us to construct a WATM network without major modification of the deployed ATM network. Most mobility support functions are performed in the peripheral network, which is a hierarchical structure constructed on top of a fixed ATM network infrastructure. This is done by applying the virtual path technique of ATM standards.

The presented location management scenario for a WATM network is based on the proposed VPWA architecture. Like most location management schemes, it consists of a location tracking scheme and a call setup scheme. The location tracking scheme makes use of multiple databases within a domain. When location changes occur within a cluster, location update messages only traverse up to the MCS. Inter-cluster location updates traverse the hierarchical tree up to the MGS of that domain. The hierarchical distributed database approach reduces the traveling distance of location tracking signals within a domain. The overhead caused by an inter-domain location tracking in IS-41 is reduced by applying forwarding pointers.

Regarding the call setup phase in the VPWA architecture, a location determination phase is performed to locate the called party before setting up a call. This is different from the standard ATM call setup procedure, which has no location query phase. The location determination process is necessary in a WATM network because the mobility of mobile hosts. The separation of ATM backbone network and the peripheral network in the VPWA architecture makes it easy to keep the location determination functions in the peripheral network. No location determination functions are needed on the ATM backbone network. After obtaining the information of the called MH's current domain, a calling MH can set up an inter-domain call across the ATM backbone network via the standard PNNI routing. The overhead incurred by call setups can be further reduced by setting forwarding pointers at selected domains. The proposed location management scheme in conjunction with the VPWA architecture can integrate a wired ATM network and a wireless network smoothly without major modifications of deployed ATM switches.

Chapter 4

Framework for WATM Handoff Management

In this chapter, we describe our WATM handoff management framework which supports seamless connection handoffs in WATM networks. The proposed handoff management scheme is based on the VPWA architecture introduced in section 3.1 of Chapter 3. This scheme requires mobility support features to be fitted into the standard ATM protocol stack and implemented in the peripheral network.

4.1 Preliminary

A handoff management mechanism on a WATM network is designed for rerouting handoff connections on the ATM backbone network so as to maintain end-to-end connections between a MH and its communication party. Several handoff management

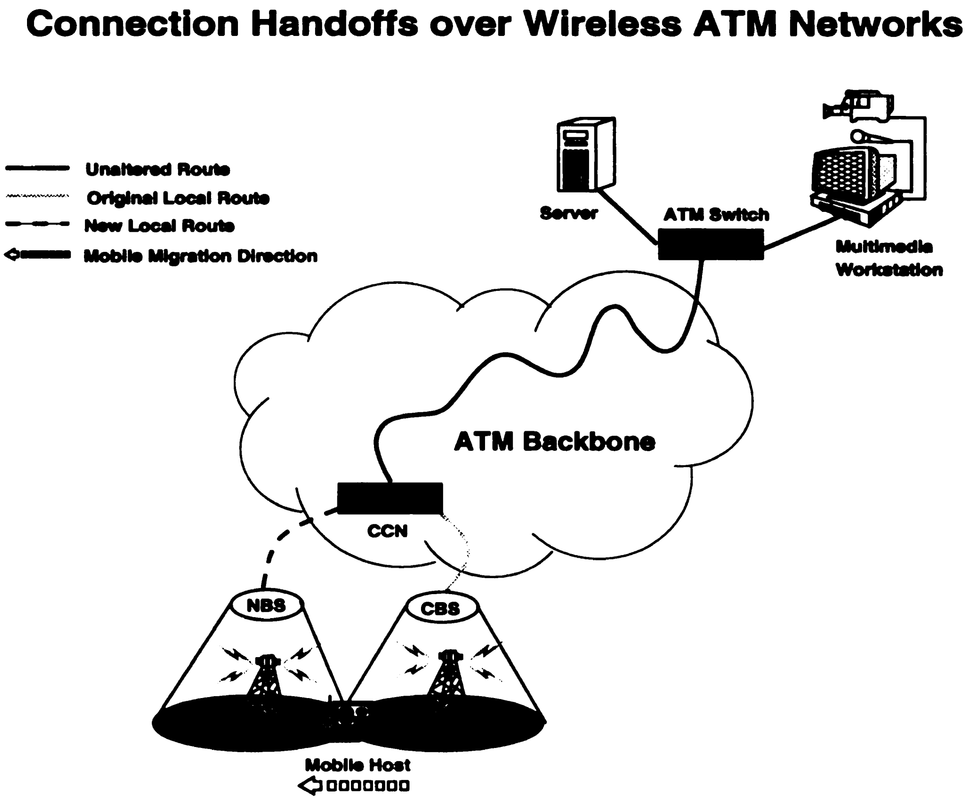


Figure 4.1: Connection handoff over a WATM network

schemes are surveyed in section 2.5.1 of Chapter 2. In the process of rebuilding handoff connections, an end-to-end re-establishment scheme is not feasible because of the tremendous connection disruption time and the network overhead. A multicast-based scheme requires multiple connections to multicast an incoming data stream to multiple BSs. This scheme consumes substantial network bandwidth and may cause the ATM cell stream to be out of sequence. An anchoring rerouting scenario can preserve cell sequencing but it suffers from scalability and inefficiency. A virtual connection tree mechanism tends to exhaust VCIs and provides no cell sequence guarantee. An incremental connection re-establishment scheme is efficient in that it can reuse an original routing path and reduce network overhead when constructing a new routing path for a handoff connection.

The central goal of an incremental connection re-establishment scheme is to reduce the connection rerouting overhead by reusing an original routing path as often as possible when constructing a new routing path for a handoff call. Recognizing that the typical movement of a mobile host is geographically local [78], most of the new routing path is common to the original one. There must be at least one common node on both the original and the new routing path. This node is called a *common cross node* (CCN). The portion shared by both the original and the new routing path is called the *unaltered route*. The routing path from the CCN to the current BS (CBS) is called the *original local route*. The *new local route* is the routing path from the CCN to the new BS (NBS). By constructing the *new local route*, joining it with the *unaltered route*, and disconnecting the *original local route*, an end-to-end connection is maintained as shown in Figure 4.1. For an incremental scheme, the closer to both base stations the CCN is the smaller the overhead of a handoff connection. The worst case occurs where the CCN is the remote host, and this reduces to an end-to-end connection re-establishment. Generally speaking, an incremental mechanism reduces both the rerouting overhead and the connection disruption time of a handoff connection.

To determine a route for a handoff connection based on an incremental mechanism, first the CCN of the new route and original route must be located. There are several CCN determining algorithms available in the literature [67, 79]. In general, these algorithms are invoked whenever a connection handoff occurs. However, a CCN determining process is a relatively time-consuming task and therefore will increase connection handoff overhead. Since the frequency of invoking this process will be very high if it is called on a per-switch handoff basis [79], it is desirable to reduce the number of invoking CCN determining processes. Considering the reduction of the CCN determination overhead, a network infrastructure is one of the major factors. A hierarchical network structure appears to be an excellent candidate for implementing an incremental connection handoff scheme due to the fact that determining a CCN is straightforward, and therefore was chosen for constructing our VPWA architecture.

4.2 WATM Handoff Management

To support seamless connection handoffs on a WATM network, we proposed a VPWA handoff management scheme [73, 74], which is based on an incremental connection re-establishment mechanism and the VPWA architecture presented in section 3.1 of Chapter 3. Following an incremental connection re-establishment discipline, the proposed VPWA handoff scheme is able to reuse the original routing path as much as possible when constructing a new routing path for a handoff connection. This reduces the network processing overhead and the connection disruption time. The hierarchical network structure of the VPWA makes the process of determining a CCN within a domain become trivial. For example, when a handoff occurs in a cluster, a MCS in Figure 3.1 is the CCN. A MGS becomes a CCN if an inter-cluster handoff occurs within a domain. Clearly, no CCN determining function is invoked by a connection handoff within a domain. Unlike the *NCNR* [67], a CCN determining process in our VPWA handoff management mechanism is incurred only when an inter-domain handoff occurs. This helps to minimize the frequency of invoking CCN determining processes.

The use of the VP technique on the VPWA network provides several advantages for supporting connection handoffs. First of all, the pre-established VP network eliminates the need for elaborate call setup functions and switching table updates along a VC route. This enables fast handoff connection setups. Second, no call admission control overhead is invoked at any intermediate nodes of a VP. Only the ends of a VP perform CAC functions. This greatly reduces the connection disruption time of a handoff connection and minimizes network resource consumption.

The proposed VPWA handoff management scheme is designed to preserve the cell sequence of a handoff connection. To preserve cell sequencing, we employ special types of cells, called Before-Handoff (BHO) cells, as a tail signal to indicate the end of transmission before a connection handoff [80]. The BHO cell is an Operation and Maintenance (OAM) cell which has the same VPI/VCI values as the user-data cells but is identified by pre-assigned code points of the Payload Type field [13]. This allows one end point of a VP, such as BSs or MCSs, to notify the other end of the status of a particular VC within a VP. Details of the process used for preserving cell sequencing is presented in section 4.5.

Regarding the design of the proposed VPWA handoff protocol, several augmented signaling messages are used to perform the handoff functions incooperated with the current ATM signaling protocol. Generally, these are implemented in the control signal layer of the ATM protocol stack. These augmented signals are listed in Table 4.1.

Abbreviation	Definition
HO_HINT	handoff hint (from MH to BS).
HO_SETUP	request for setting up new VC for handoff connection.
HO_CONN	confirmation of setting up new VC for handoff connection.
HO_REQ_SETUP	request for handoff to NBS (from MH to BS).
HO_REQ_CONN	confirmation of handoff to NBS (from BS to MH).
HO_EXT_SETUP	request for setting up extension for handoff connection (from CBS to NBS).
HO_EXT_CONN	confirmation of a connection extension request (from NBS to CBS.
HO_FWD_SETUP	request for setting up a forward connection (from NBS to CBS).
HO_FWD_CONN	confirmation of setting up a forwarding connection (from CBS to NBS).
HO_MIG_IND	indication of a MH migration (from CBS to NBS).
HO_MIG_NTF	notification of a MH migration.
HO_RROUT_IND	indication of rerouting a virtual connection.
HO_RROUT_SETUP	request for setting up a rerouting virtual connection.
HO_RROUT_CONN	confirmation of setting up a rerouting virtual connection.

Table 4.1: Augmented signaling message for supporting handoff

Before we discuss our VPWA handoff scheme for various handoff conditions, a few preliminary conditions must be described. To assist in a handoff process, each domain of a VPWA network is assigned a particular segment of ATM addresses. This is supported by the ATM PNNI standard. Each BS has the addresses of all its neighboring BSs'. Like most wireless systems, every BS in the VPWA architecture periodically broadcasts a beacon signal which contains the identification (e.g., ATM address) of the BS within its radio coverage.

Notation	Definition
BW_{vci}	Bandwidth of the handoff VCI connection.
N_a	Number of hops from CBS to its MCS and vice versa.
N_{ax}	Number of hops from a MCS in domain A to *MGS A* and vice versa.
N_b	Number of hops from NBS to its MCS and via versa.
N_{bx}	Number of hops from a MCS in domain B to *MGS B* and via versa.
N_{cx}	Number of hops from *MGS A* to CCN and via versa.
N_{dx}	Number of hops from *MGS B* to CCN and via versa.
PD	Propagation delay of wireline link
PD_{wl}	Propagation delay of wireless link
PPT_{cac}	Protocol processing time for steps where admission control is to be performed, in excess of fixed protocol processing time.
PPT_{cwa}	ATM layer and wireless sublayer conversion time.
PPT_x	Protocol processing time at x, x can be MH, BS, MCS, MGS, or CCN.
RT_k	Relative time that message k has reached its destination and been processed. Assume that the time when a MH decides to perform handoff is the reference time 0.
T_{acq}	Time for a MH to acquire a channel in a wireless cell.
T_{sw}	Time for a *CCN* to switch a connection from the *original local route* to the *new local route*.
T_{BHO}	Time when a BHO cell is issued until it reaches its destination.
T_{n1}	Time when the first data cell traverses along the *new local route* and arrives at NBS or CCN.
TD_c	Transmission time for sending a control packet over a control channel at any wireline ATM node.
TD_d	Transmission time for sending an ATM cell at any wireline ATM node.
TD_{wlc}	Transmission time for sending a control wireless sublayer packet at MH or BS.
TD_{wld}	Transmission time for sending a data wireless sublayer packet at MH or BS.

Table 4.2: Definitions of parameters and notation

To evaluate a connection handoff, the connection disruption time and the buffer requirements are two major metrics. The connection disruption time is the time during which a handoff call loses its connection during a handoff process. It is desirable to minimize this time. On the other hand, extra buffer space may be needed when a handoff connection is rerouted during a handoff session. The demanded buffer space should also be minimized to reduce the system overhead. We will analyze our VPWA handoff protocol based on these two performance metrics. Notation and parameters to be used in the analyses of the proposed handoff protocol are defined in Table 4.2.

Upon examining various handoff cases of applying our VPWA handoff scheme, we discuss one handoff connection only to simplify the description of the handoff algorithm. Multiple handoff connections are supported by the scheme. According to the arrangement of BSs and the use of wireless technologies, a connection handoff can be classified as a *handoff with-hint* and a *handoff without-hint.* A *handoff with-hint* means that a MH can inform the CBS of its departure during a handoff session. Otherwise, it is called a *handoff without-hint.* Depending on the traffic flow direction of a handoff connection, a handoff connection can be divided into: (1) an *uplink* handoff if traffic flows from a MH to the wired backbone network, (2) or a *downlink* transmission if traffic flows from a wired backbone network to a MH. Also according to the VPWA architecture a connection handoff can be either an intra-domain or an inter-domain handoff. Based on the above classifications, we discuss various handoff cases in the following sections.

4.3 Intra-Domain Connection Handoff Management

In this section, we present our connection handoff schemes for handling both uplink and downlink intra-domain connection handoffs on a VPWA network. Considering a domain in the VPWA network, intra-domain connection handoffs can also be classified into intra-cluster (intra-domain) handoffs and inter-cluster (intra-domain) handoffs. The procedure of an intra-cluster handoff is simpler than the procedure of an inter-cluster handoff. The signaling message flows of these two procedures are very similar, except there is no signal traversing up to the MGS in intra-cluster handoffs. To avoid redundancy, we omit the description of every signaling step of an intra-cluster handoff procedure. Signaling message flows of various intra-cluster connection handoff cases are shown in Figure 4.2 (a), (b), (c), and (d) for: uplink handoff with-hint; downlink handoff with-hint; uplink handoff without-hint; downlink handoff without-hint; respectively. The derivation of system parameters (connection disruption time and buffer requirement) for each of the handoff conditions can be found in Appendices B.1– B.4.

4.3.1 Intra-Domain Uplink Connection Handoff

A simplified procedure for an inter-cluster (intra-domain) uplink connection handoff with-hint is shown in Figure 4.3. The timeline of the signal message flow is illustrated in Figure 4.4. It should be noted that the message number has no absolute relation

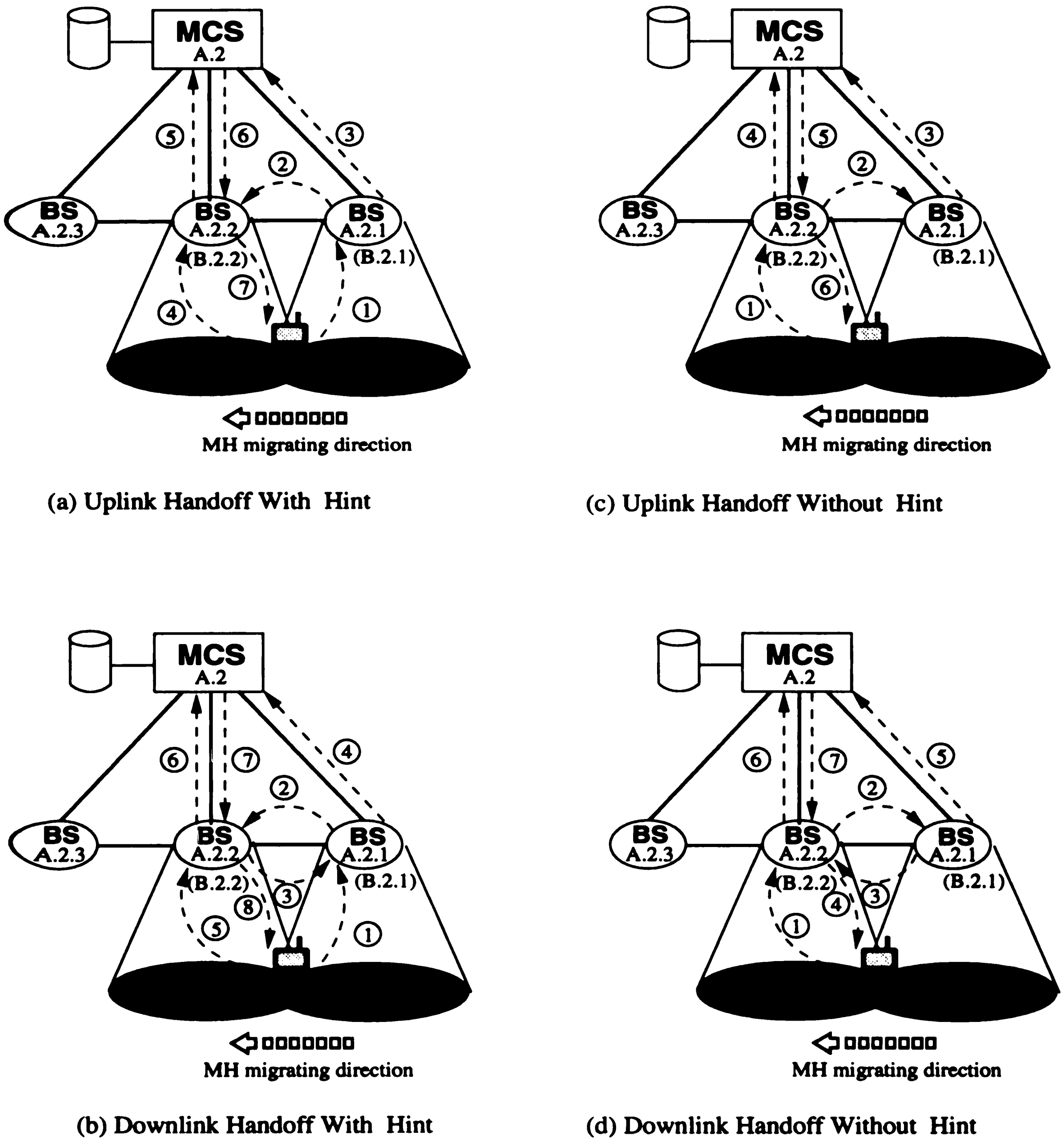


Figure 4.2: WATM Inter-Domain connection handoff (Phase 1)

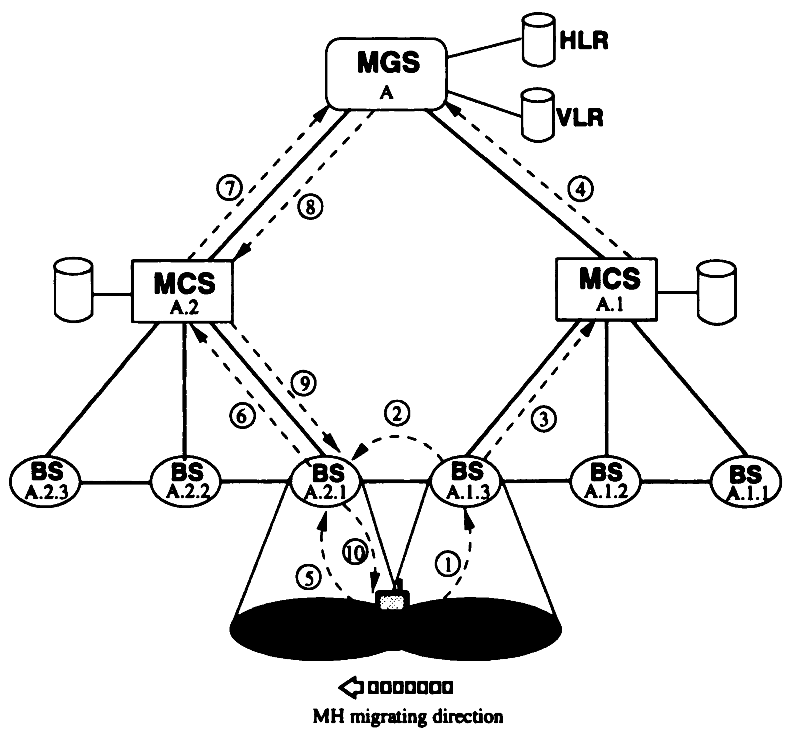


Figure 4.3: VPWA procedure for an inter-cluster uplink handoff with-hint

with the order of the message. The description for each message in the procedure for an inter-cluster (intra-domain) uplink connection handoff with-hint is listed below:

1. When a MH moves into the handoff zone between the CBS (*BS A.1.3*) and the NBS (*BS A.2.1*), it detects the existence of *BS A.2.1*, via capturing the emitted beacon signal. It sends a signal HO_HINT to inform *BS A.1.3* of its migration to the new BS. This signal contains the address of *BS A.2.1*.

2. Once *BS A.1.3* has received the signal HO_HINT, it sends a signal message (HO_MIG_IND) to the NBS (*BS A.2.1*) to indicate the arrival of the MH.

3. *BS A.1.3* also sends the signal HO_MIG_NTF to its MCS (*MCS A.1*). This message is used to inform the MCS of the MH's migration. Upon receiving

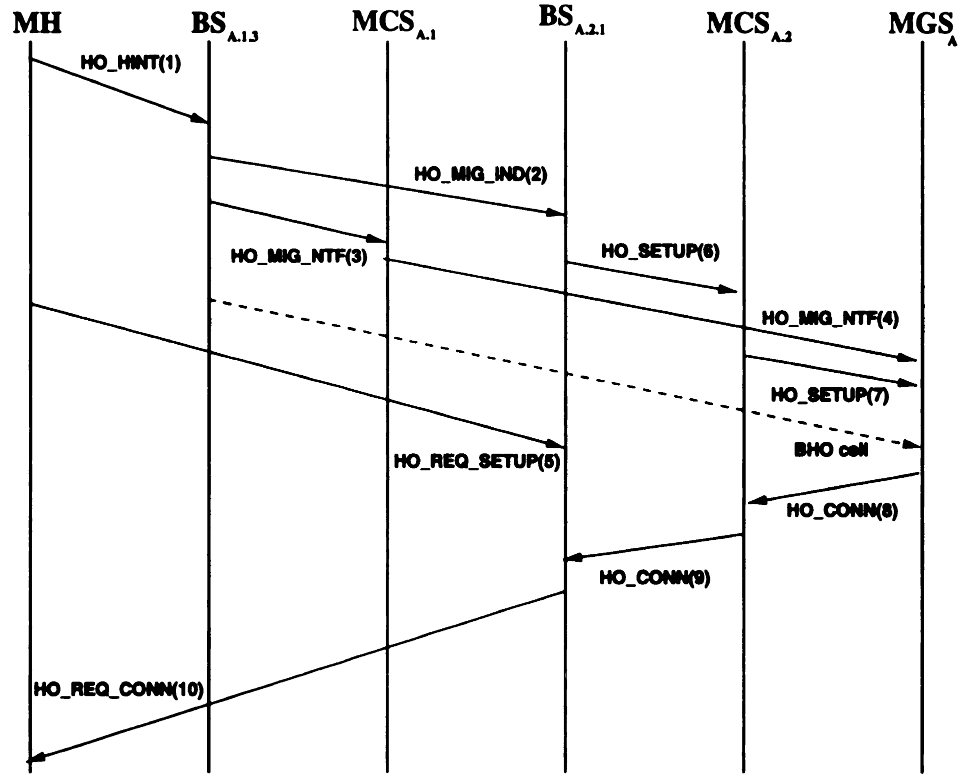


Figure 4.4: Time line of VPWA procedure for an inter-cluster uplink handoff with-hint

this signal, *MCS A.1* marks its local database to indicate the departure. It then forwards this signal to its MGS (*MGS A*). At the same time, *BS A.1.3* composes and sends a BHO cell along the uplink *original local route* of the handoff call.

4. When *MGS A* receives the signal HO_MIG_NTF, it marks the database entry of the MH in its VLR or HLR.

5. Immediately after the MH sends out the HO_HINT, it attempts to lock onto a control channel in the new wireless cell of *BS A.2.1.* Once it obtains the control channel, it sends a message HO_REQ_SETUP to *BS A.2.1.* This signal message contains the tuple (TAN, PAN, UID) of the MH and the address of the CBS

(*BS A.1.3*).

6. When *BS A.2.1* receives either a signal HO_MIG_IND or a HO_REQ_SETUP, whichever arrives at *BS A.2.1* first, it composes and transmits a message HO_SETUP to its MCS (*MCS A.2*). This signal is to setup a uplink *new local route* for the handoff call.

7. When *MCS A.2* receives the HO_SETUP, it assigns a TAN address for the MH and records the tuple (TAN, PAN, UID) of the MH in its local database. It modifies and forwards the request signal HO_SETUP to *MGS A*.

8. Upon receiving the signal HO_SETUP, *MGS A* performs authentication by examining information in this signal as well as the signal HO_MIG_NTF. When the authentication is passed, *MGS A* updates the information of the MH in its VLR or HLR database and marks *MCS A.2* as the serving MCS. It also performs a simple CAC function and sends a connection acknowledgment (HO_CONN) back to *MCS A.2*.

9. After receiving the message HO_CONN, *MCS A.2* forwards the signal to *BS A.2.1*. The new TAN address of the MH may be included in this message; therefore, no additional messages for delivering the new TAN address to the MH are necessary. When *BS A.2.1* receives the signal, a *new local route* for the handoff call between *BS A.2.1* and *MGS A* is established.

10. When *BS A.2.1* receives the signal HO_CONN, it sends the signal

HO_REQ_CONN to the MH as a response to the signal HO_REQ_SETUP sent by the MH. Once the MH receives the message HO_REQ_CONN, the handoff call has an end-to-end connection again. It can now resume its uplink transmission.

To inform the CCN (*MGS A*) when to switch the handoff connection from the *original local route* to the *new local route*, an in-band BHO cell is sent along the original routing path by *BS A.1.3*. Once *MGS A* detects the BHO cell, it can join the *unaltered route* of the VC with the *new local route* from *BS A.2.1*. However, extra care needs to be taken when applying this scheme to realtime traffic. Although it must be very rare that the first cell from *BS A.2.1* would arrive at *MGS A* and times out in the queue before the arrival of the BHO cell, we handle that situation by switching from the *original local route* and the *new local route*, triggered by either the arrival of the BHO cell or the time stamp of the first cell in the queue. Generally, a BHO cell should arrive at *MGS A* before the arrival of the first ATM cell from the *new local route*. Thus, no buffer space is needed to buffer cells from the *new local route*, and in this case cells from the *new local route* arrives at *MGS A* before the arrival of the BHO cell, these cells must be buffered. As derived from Appendix B.5, the connection disruption time ($T_{disrupt}$) and buffer requirements at *MGS A* are:

$$
\begin{aligned}
Tdisrupt &= RT_{10} \\
&= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 2PPT_{cac} + 2(N_b + N_{bx})(TD_c + PD + PPT_x) \\
&\quad + min\{N_h(TD_c + PD + PPT_x), (T_{acq} + PPT_{cac})\} \\
B_{MGS} &= max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\} BW_{vci}
\end{aligned}
$$

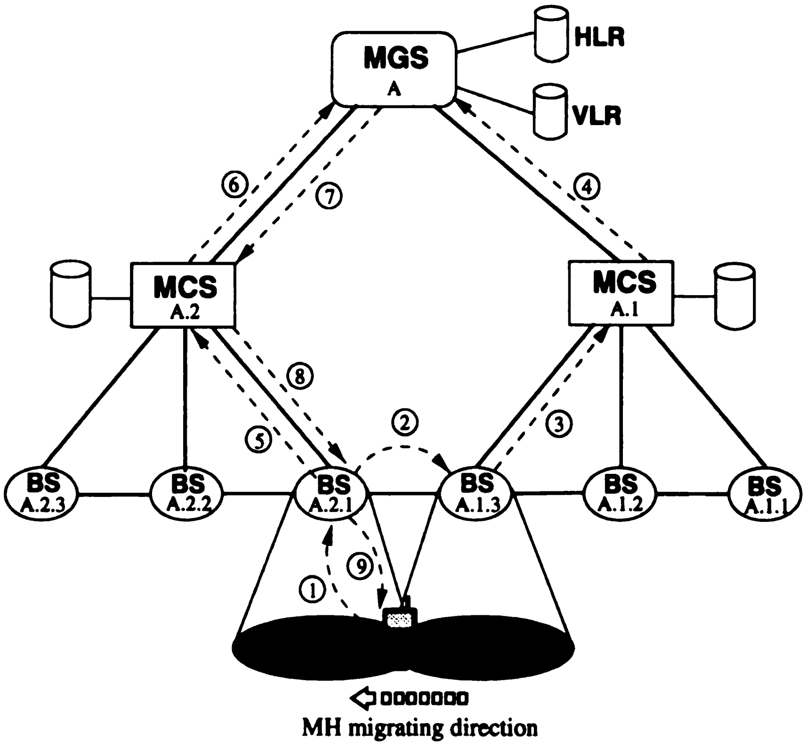


Figure 4.5: VPWA procedure for an inter-cluster uplink handoff without-hint

$$= max\{[T_{sw} + (TD_D + PD)(N_a + N_{ax} - N_b - N_{bx}) - (4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_{cac})$$
$$-2(N_b + N_{bx})(TD_c + PD + PPT_x) - min\{N_h(TD_c + PD + PPT_x), (T_{acq} + PPT_{cac})\}], 0\}BW_{vci}$$

However, there would be no buffer requirements at the CCN (*MGS A*), since the BHO cell usually arrives at the MGS much earlier than the arrival of cells from the *new local route*. That is, when cells along the *new local route* arrive at *MGS A*, the *unaltered route* has already been switched from the *original local route* to the *new local route*.

In the case that a handoff hint is not possible when a MH handoffs from one BS to another, Figure 4.5 illustrates the signal message flow that is performed during an inter-cluster (intra-domain) uplink handoff without-hint. The scheme is very similar to the with-hint scheme that is presented above. Comparing Figure 4.3 with Figure 4.5, the only difference between them is that the handoff procedure is triggered by the HO_HINT signal in the former and by the HO_REQ_SETUP signal in the later. The latter incurs a slightly longer delay in completing the handoff process because it has to wait until the new BS receives the HO_REQ_SETUP signal. Actually, the latter can be used to complement the former in the case where the HO_HINT signal is lost during the handoff of a MH. As is derived in Appendix B.6, the disruption time and the buffer size needed at *MGS A* are:

$$
\begin{aligned}
Tdisrupt &= RT_9 \\
&= T_{acq} + 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 3PPT_{cac} + 2(N_b + N_{bx})(TD_c + PD + PPT_x) \\
B_{MGS} &= max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\}BW_{vci} \\
&= max\{[T_{sw} + (TD_d + PD)(N_a + N_{ax} - N_b - N_{bx}) - (TD_c + PD + PPT_x)(N_h - 2N_b - 2N_{bx}) \\
&\quad -(2PPT_{cwa} + TD_{wlc} + PD_{wl})], 0\}BW_{vci}
\end{aligned}
$$

Again, similar to the uplink with-hint case, generally no buffer space is needed to buffer cells from the *new local route* at the CCN (*MGS A*).

4.3.2 Intra-Domain Downlink Connection Handoff

Figures 4.6 and 4.7 show the signaling procedure and the signaling timeline of an inter-cluster and intra-domain downlink handoff, respectively. The steps of the procedure are described below:

1. When the MH enters the handoff zone, it picks up the beacon signal transmitted

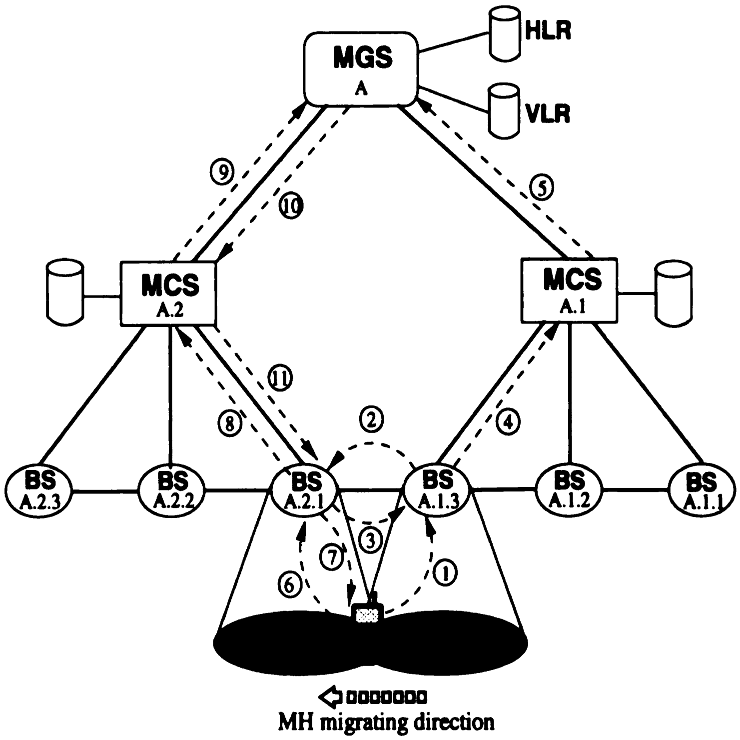


Figure 4.6: VPWA procedure for an inter-cluster downlink handoff with-hint

by the NBS (*BS A.2.1*). The MH sends the HO_HINT to the CBS (*BS A.1.3*) to indicate its migration and the NBS (*BS A.2.1*).

2. *BS A.1.3* sends the signal HO_EXT_SETUP to *BS A.2.1* when it receives the HO_HINT. A HO_EXT_SETUP is a request to setup an extended VC connection between *BS A.1.3* and *BS A.2.1* for the handoff call associated with the MH.

3. *BS A.2.1* acknowledges the request, HO_EXT_SETUP, by sending a response signal HO_EXT_CONN back to *BS A.1.3*. When *BS A.1.3* receives this signal, the ***original local route*** of the handoff call is extended to *BS A.2.1*. A downlink ATM cell stream of the handoff call is then forwarded to *BS A.2.1*.

4. After receiving the HO_HINT, *BS A.1.3* also sends out the signal HO_MIG_NTF

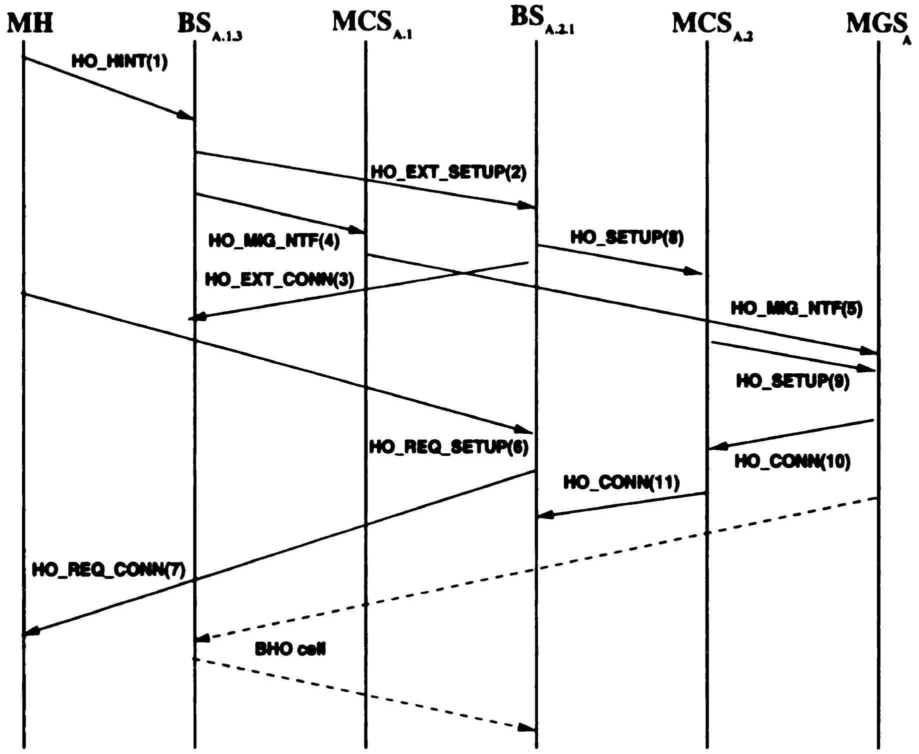


Figure 4.7: Time line of VPWA procedure for an inter-cluster downlink handoff with-hint

to its MCS (*MCS A.1*). This signal is used to notify *MCS A.1* about the migration of the MH. Once *MCS A.1* receives the signal HO_MIG_NTF, it records the MH's departure in its local database and forwards this signal to *MGS A*.

5. Upon receiving the signal HO_MIG_NTF, *MGS A* marks the database entry of the MH at either the VLR database or the HLR database.

6. Immediately after the MH sends out a HO_HINT, it searches the control channel in the new wireless cell of *BS A.2.1*. Once it locks onto the control channel, it sends the message HO_REQ_SETUP to *BS A.2.1*. This signal message contains the tuple (TAN, PAN, UID) of the MH and the address of the CBS (*BS A.1.3*),

and is used for registering the MH at *MCS A.2*.

7. *BS A.2.1* sends a HO_REQ_CONN back to the MH when it either sends out a HO_EXT_CONN or receives a HO_CONN from *MCS A.2*. This signal is to inform the MH that an end-to-end connection for the handoff call is established.

8. When receiving either a HO_EXT_SETUP or a HO_REQ_SETUP, *BS A.2.1* sends signal HO_SETUP to its MCS (*MCS A.2*) to setup a downlink connection for the handoff call. This signal includes the tuple (TAN, PAN, UID) of the MH.

9. After *MCS A.2* receives a HO_SETUP, it assigns a new TAN address for the MH. This new TAN address replaces the TAN address in the tuple before it is recorded in the local database. *MCS A.2* forwards the HO_SETUP to *MGS A*.

10. *MGS A* honors the connection setup request in the signal HO_SETUP by sending the signal HO_CONN back to *MCS A.2*. It also performs an authentication check when it receives both the HO_SETUP and the HO_MIG_NTF. If authentication is successful, *MGS A* detaches the *original local route* from the *unaltered route* and joins the *unaltered route* with the *new local route*. After that, it sends a BHO cell along the *original local route* to *BS A.1.3*.

11. *MCS A.2* forwards the HO_CONN signal to *BS A.2.1* when it receives the signal.

12. Once *BS A.2.1* receives the HO_CONN, a *new local route* for the handoff connection is established.

In the downlink case, the original routing path of the handoff call is temporarily extended along the pre-setup VP from the CBS (*BS A.1.3*) to the NBS (*BS A.2.1*). Cells arriving at the CBS when the MH is absent are forwarded to the NBS and later delivered to the MH. In the meantime, the CCN (*MGS A*) sends a BHO cell along the *original local route*. This BHO cell eventually arrives at NBS (*BS A.2.1*) and triggers the NBS to switch the VC connection from the temporarily extended route to the *new local route*. Considering the case of time sensitive traffic, occasionally, the path extension scheme may deliver outdated cells to the MH since they traverse a longer path. To prevent unwanted delivery, the CBS decides whether to extend the path or not by periodically collecting the round trip time information for all its neighboring BSs. This is made possible by sending resource management (RM) cells on the VPs between neighboring BSs [13].

As derived from Appendix B.7, the connection disruption time ($T_{disrupt}$) and the buffer requirements on the CBS (*BS A.1.3*) and on the NBS (*BS A.2.1*) are:

$$
\begin{aligned}
T_{disrupt} &= RT_7 \\
&= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + min\{N_h(TD_c + PD + PPT_x), T_{acq} + PPT_{cac} + 2(N_b \\
&\quad + N_{bx})(TD_c + PD + PPT_x)\} \\
B_{NBS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci} \\
&= [(2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 2PPT_{cac}) + 2N_h(TD_c + PD + PPT_x)]BW_{vci}
\end{aligned}
$$

Figure 4.8 shows the signal procedure for an inter-cluster, intra-domain downlink handoff when a handoff hint is not possible. As in the uplink case, this scheme can be used to complement the with-hint scheme so as to provide a smooth downlink

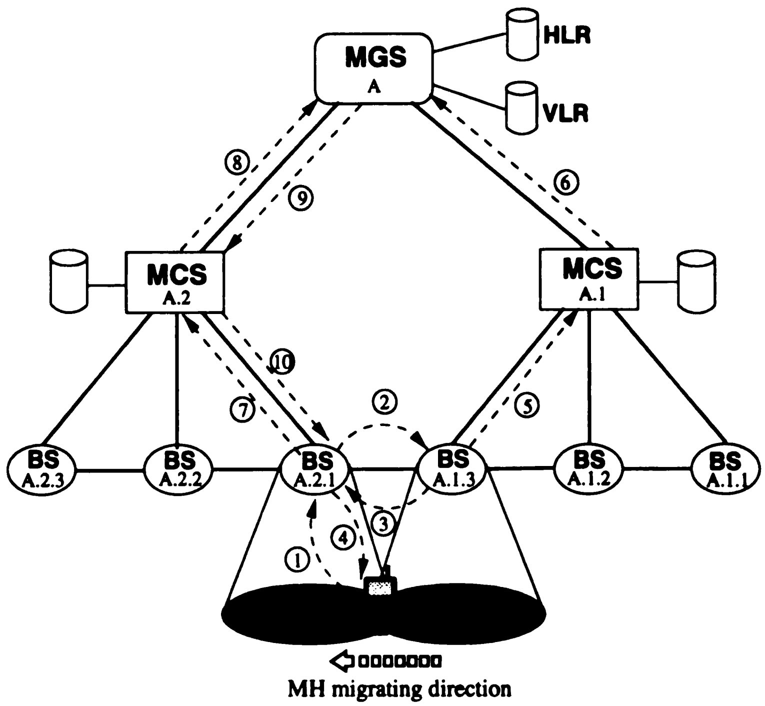


Figure 4.8: VPWA procedure for an inter-cluster downlink handoff without-hint

handoff process. As derived from Appendix B.8, the connection disruption time, and the buffer requirements on the CBS and on the NBS are:

$$
\begin{aligned}
T_{disrupt} &= RT_4 \\
&= T_{acq} + 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 3PPT_{cac} + 2N_h(TD_c + PD + PPT_x) \\
B_{CBS} &= [RT_2 + 2(TD_c + PD + PPT_x)]BW_{vci} \\
&= [2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 3PPT_{cac} + (N_h + 2)(TD_c + PD + PPT_x)]BW_{vci} \\
B_{NBS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci} \\
&= max\{(TD_d + PD)(N_{ax} + N_a + N_h - N_{bx} - N_b) + PPT_x + T_{sw}, 0\}BW_{vci}
\end{aligned}
$$

Note that most signal messages from the location tracking scheme and from the handoff scheme can be merged in all intra-domain handoff cases. This should greatly reduce the number of signals exchanged on the network and minimize the signal overhead. Take an inter-cluster downlink handoff case as an example. Figures 3.3 and 4.6 show the signal diagram of the location tracking scheme and the handoff rerouting scheme, respectively. Signals (1)–(5) in Figure 3.3 can be merged with their counterpart signals (6), (8), (9), (10), and (11) in Figure 4.6, respectively.

4.4 Inter-Domain Connection Handoff Management

When a MH moves from one domain to another domain, an inter-domain connection handoff procedure is needed to maintain end-to-end connections for handoff calls associated with the MH. Generally, an inter-domain connection handoff procedure is more complicated and incurs more overhead. If MHs continually move back and forth between neighboring domains, they may cause prohibitive network overhead and possibly exhaust the whole system. As described in the design of the VPWA architecture in section 3.1 of Chapter 3, a special arrangement of *border MCSs* is used to relieve this thrashing problem. The use of *border MCSs* to relieve the thrashing problem on location management has been presented in section 3.2 of Chapter 3. The relief of the thrashing phenomenon in handoff management is addressed in this section.

According to the VPWA architecture, a MH is handed over from one domain to another. The process of an inter-domain handoff (with or without hint) can be divided into two phases. Detailed operations of each phase are presented as follows.

Phase 1

When a MH moves from *BS A.2.1* to P*BS A.2.2*, an intra-cluster connection handoff procedure is invoked. The signaling procedure of this phase is exactly the same as the procedure of an intra-cluster, intra-domain connection handoff. Signaling flow diagrams of various handoff conditions were presented in Figure 4.2. After completing this phase, the MH still belongs to domain A. In this phase, system parameters, such as the connection disruption time and the buffer requirements on the CBS and the NBS, are the same as those of intra-cluster handoffs. Derivations of these parameters for various handoff conditions can be found in Appendices B.1– B.4.

Phase 2

After the MH is settled down in the wireless cell of *BS A.2.2*, *MCS A.2* detects that the MH has moved deeply into the coverage of the cluster and is approaching domain B. It then decides to migrate the MH from domain A to domain B. There are two major tasks in this phase. First, the registration of the MH must be migrated from domain A to domain B. Second, connections must be rerouted from domain A to domain B. The procedure of migrating the registration of the MH is very similar to the procedure of an inter-domain location tracking illustrated in Figure 3.4, except that the registration request is issued by *MCS A.2* instead of the MH. The signaling message flow of this procedure is shown in Figure 4.9. When this registration migration procedure is done, the MH belongs to domain B. Now, the MH maintains the TAN address newly assigned by the *new visiting MGS* (*MGS B*) in domain B. Also, both the local database and the VLR database indicate that the MH is in domain B by recording the tuple (TAN, PAN, UID) of the MH. A forwarding pointer, pointing to *MGS B*, is set in the HLR of the *home domain*. The database entry of the MH in *MGS A* is now pointing to *MGS B*. Thus, if a subsequent call for the MH arrives at *MGS A*, it is forwarded to *MGS B* according to this pointer.

The signaling message flow for rerouting connections is shown in Figure 4.10. The detailed procedure of this phase is listed in the following steps:

1. *MCS A.2* sends the signal HO_RROUT_SETUP to *MGS B* for setting up a new routing path in the domain B for the handoff call.

2. *MCS A.2* also sends the signal HO_RROUT_REQ to *MGS A* for rerouting handoff connections.

3. MSG A determines the *common cross node* of the handoff call via invoking a CCN determining process. Once it locates the CCN, it forwards the signal HO_RROUT_REQ to the CCN.

4. MSG B calls the same CCN determining function to locate the CCN. After that, it forwards the HO_RROUT_SETUP to the CCN.

5. When the CCN receives both the HO_RROUT_SETUP and the HO_RROUT_REQ, it sends the signal HO_RROUT_CONN back to *MGS B*. If the handoff call is a downlink connection, the CCN sends a BHO cell along the *original local route* of the handoff call to mark the end of the downlink cell stream along this route.

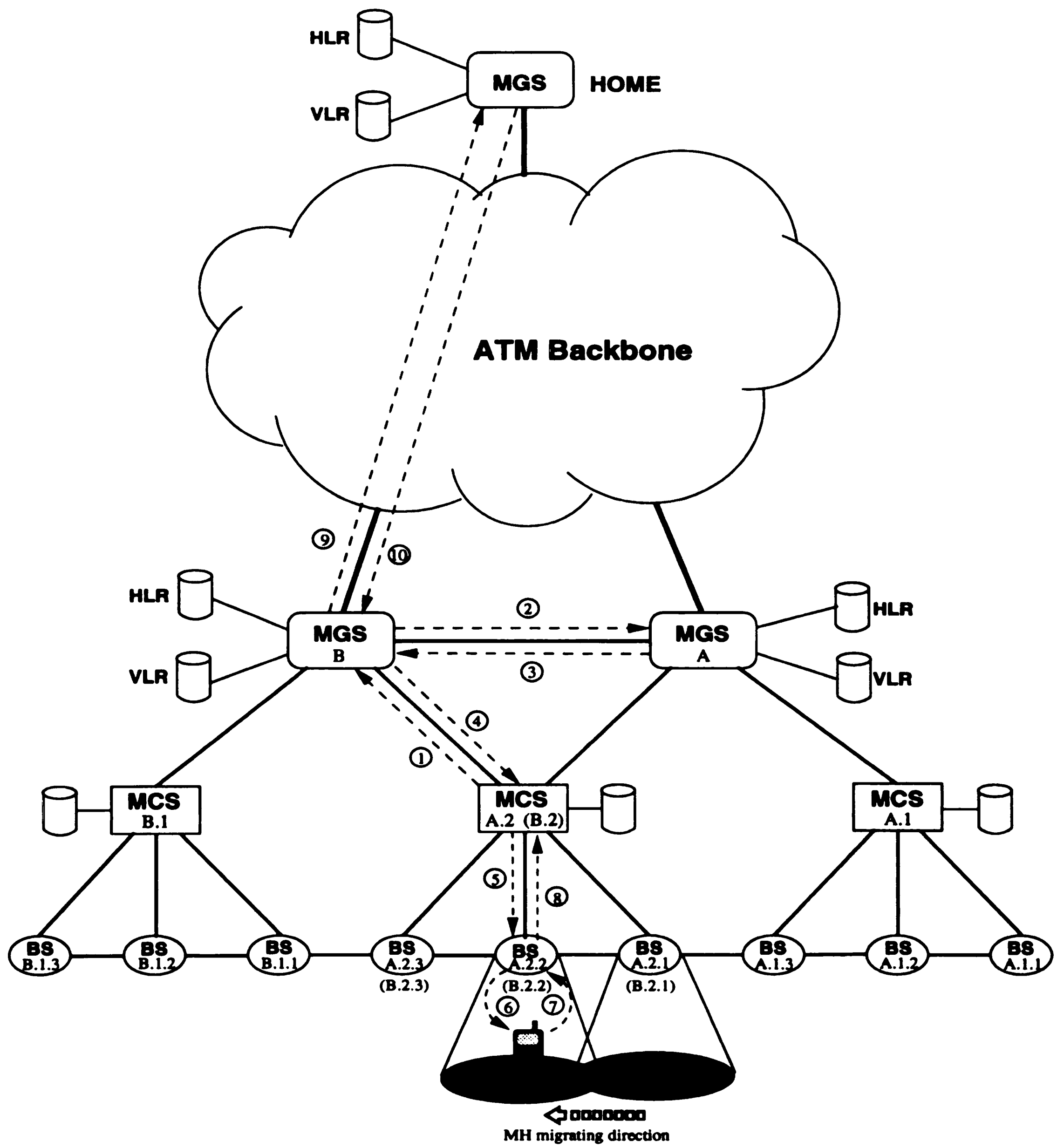


Figure 4.9: WATM Inter-Domain handoff (Phase 2): registration migration

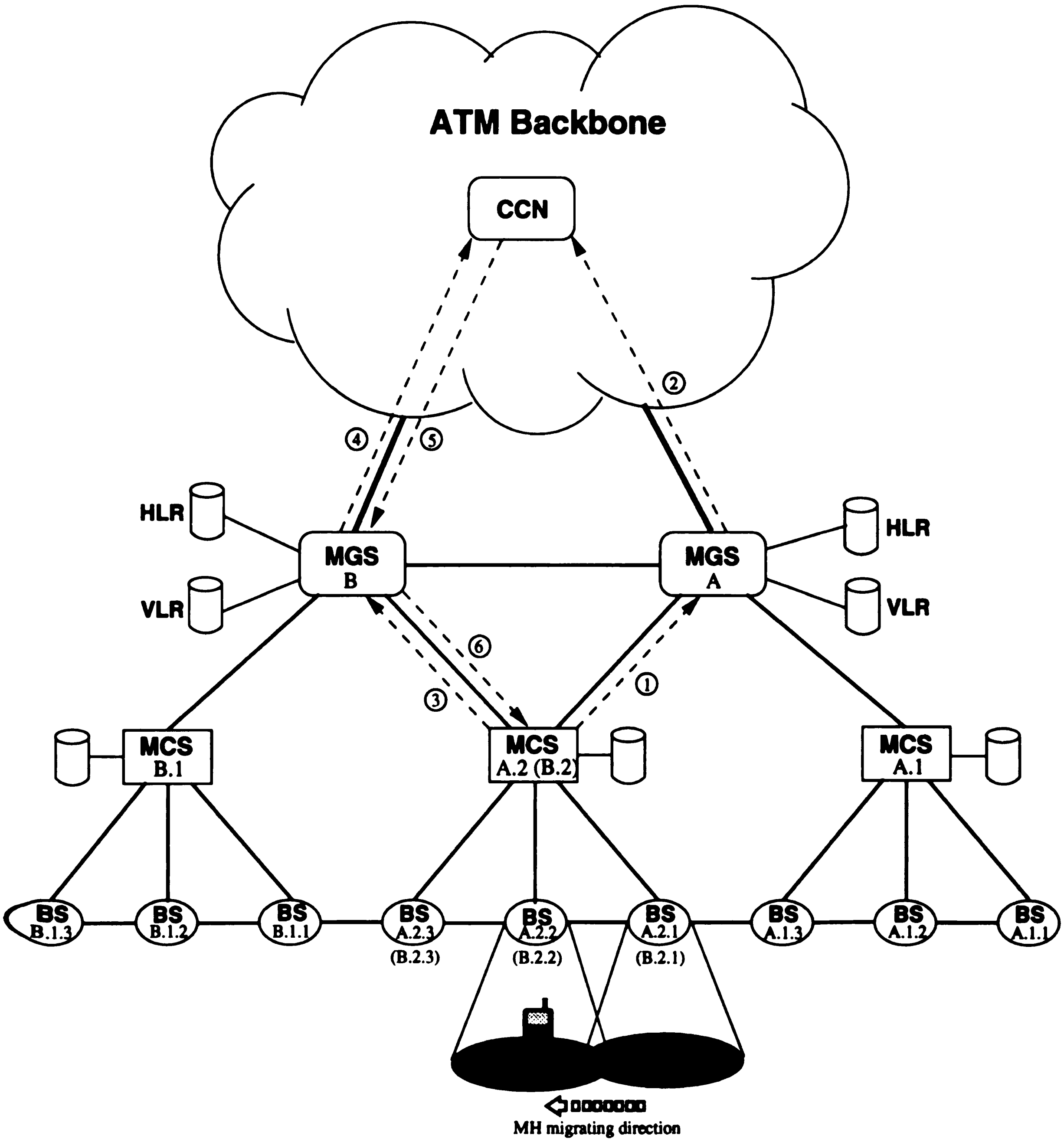


Figure 4.10: WATM Inter-Domain handoff (Phase 2): connection rerouting

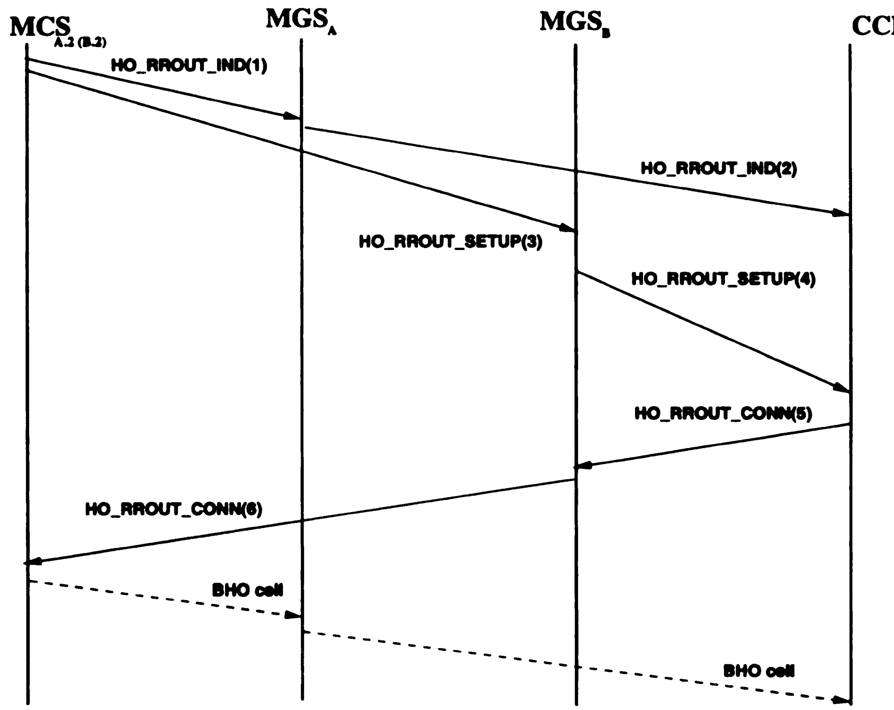


Figure 4.11: Time line of the VPWA procedure for an inter-domain uplink handoff in Phase 2

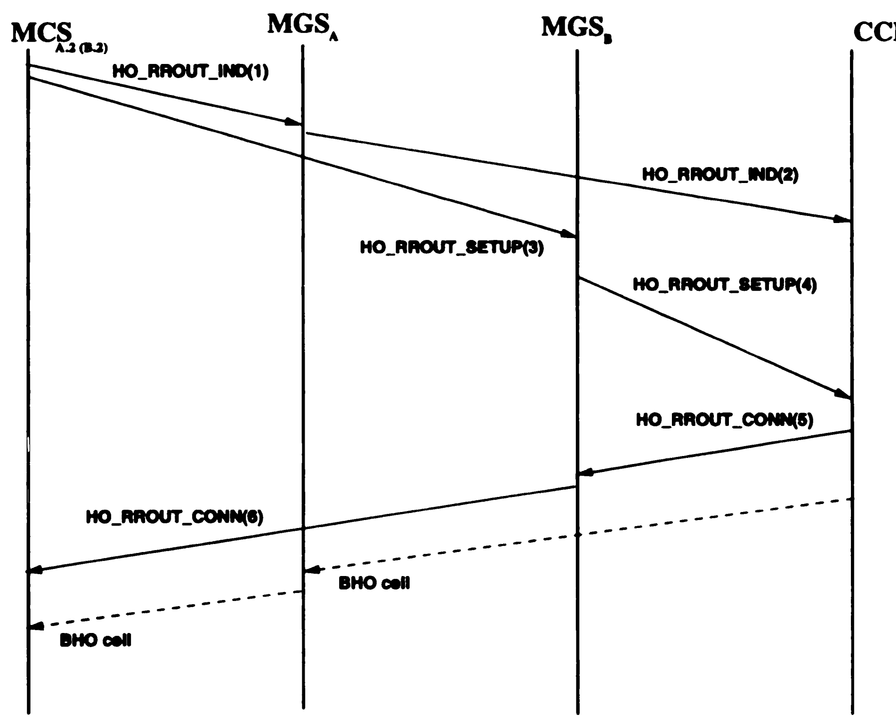


Figure 4.12: Time line of the VPWA procedure for an inter-domain downlink handoff in Phase 2

6. *MGS A* forwards the HO_RROUT_CONN to *MCS B.2 (A.2)* when it receives the signal.

7. Once *MCS B.2 (A.2)* receives the HO_RROUT_CONN, a new routing path between the CCN and *MCS B.2* is established. If the handoff call is an uplink connection, the *MCS B.2 (A.2)* sends a BHO cell along the original routing path to mark the end of the cell stream along this route.

To complete a handoff process, extra steps must be taken for both uplink handoff and downlink handoff calls. Time line diagrams of signaling flows for inter-domain uplink and downlink handoffs in phase 2 are shown in Figure 4.11 and Figure 4.12, respectively. In a uplink handoff case, when the CCN receives the BHO cell sent by *MCS A.2 (B.2)* along the *original local route*, it knows that this is the end of the cell stream along that route. Therefore, the CCN detaches the *original local route* from the *unaltered route* and joins together the *new local route* with the *unaltered route.* It sends a **CONNECT RELEASE** signal to *MCS A.2* to tear down the *original local route.* Until now, the uplink handoff process is completed. In the case of a downlink handoff process, when *MCS B.2* detects the BHO cell sent by the CCN, it detaches the *original local route* from the route to *BS B.2.2.* The route between *MCS B.2* and *BS B.2.2* is joined to the *new local route* (which is between *MCS B.2* and the CCN). This completes the downlink handoff process.

Before a *border MCS* initiates the process of this phase, the CCN on the ATM backbone must be determined. Since the CCN determining process is known to be very time-consuming, invoking this function in a handoff session is not desirable. By separating an inter-domain handoff into two phases, we are able to avoid calling a CCN determining process during a connection handoff. Therefore, the completion time of this phase will not affect the connection disruption time of a handoff call on the MH. This allows the connection disruption time of an inter-domain handoff to be remained the same as that of an intra-cluster handoff (invoked in phase 1). Buffer space is needed at either *MCS A.2* or the CCN, depending on the direction of a handoff call. The required buffer spaces for uplink handoffs and downlink handoffs (derived from Appendix B.9.1 and Appendix B.9.2) are listed below:

Uplink

$$\begin{aligned} B_{CCN} &= max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\}BW_{vci} \\ &= max\{(TD_d + PD)(N_{ax} + N_{cx} - N_{bx} - N_{dx}) + PPT_x + T_{sw}, 0\}BW_{vci} \end{aligned}$$

Downlink

$$\begin{aligned} B_{MCS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci} \\ &= max\{(TD_d + PD)(N_{ax} + N_{cx} - N_{bx} - N_{dx}) + PPT_x + T_{sw}, 0\}BW_{vci} \end{aligned}$$

Note that similar to intra-domain handoff cases, some signals of the location management scheme and the handoff management scheme can be combined to further reduce the number of signals on the network. For example, the signal message (1) in Figure 4.9 can be merged with the signal message (3) in Figure 4.10. This reduces the signaling overhead for inter-domain handoffs.

4.5 Preserving Cell Sequencing

In wireline ATM networks, keeping cells in sequence is not a problem since cells traverse along the same connection in First-In-First-Out (FIFO) fashion for the whole communication session. However, in a wireless environment, as a mobile host roams among wireless cells, the routing path needs to be updated with the changing location of the mobile host. In packet switching networks, this is not a problem since packets are routed to their destinations. The higher protocol layers, such as TCP, at the receiving end takes care of packet sequencing. But in wireless ATM network, ATM cells flow along a pre-setup circuit. The ATM layer is not responsible for cell sequencing. Therefore, cells routed to the CBS after the MH moves will be either discarded or forwarded. Discarding cells is undesirable since it degrades the QoS. Cells forwarded to the NBS may arrive at the NBS later than cells from the new routing path. This causes the cell stream of a handoff call to be out-of-sequence. At higher network protocol layers, packets containing out-of-sequence cells are considered errors and will be discarded. For loss sensitive applications, this requires end-to-end retransmission [13]. End-to-end retransmission demands substantial buffer space to accommodate incoming cells while retransmission in proceeding. Also, it consumes extra network resources, such as network bandwidth and switch buffers. The problem could cause a network catastrophe since retransmitted packets might be lost and require retransmissions again. Therefore, the preservation of a connection's cell sequence is considered essential on wireless ATM networks.

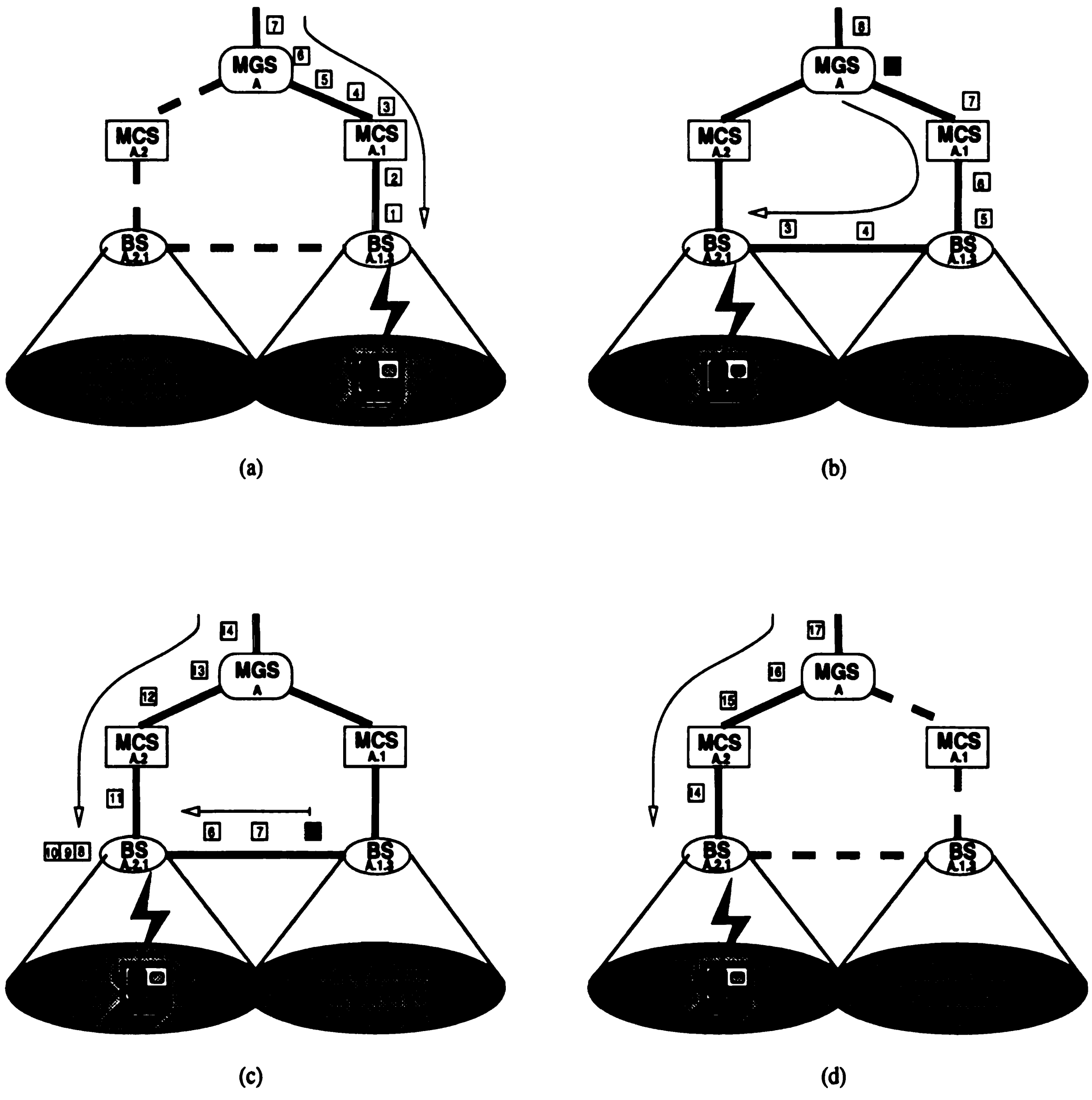


Figure 4.13: The VPWA handoff scheme for preserving cell sequence in an inter-cluster downlink handoff case

Our VPWA handoff schemes presented in the previous sections can preserve cell sequencing in handoff sessions. We use a BHO cell as an indication of the end of the cell stream traversing along the *original local route.* The merging node of the new routing path and the original routing path can only deliver cells from the *new local route* if it detects a BHO cell. By doing this, the cell sequence of a handoff call on the ATM network is preserved. Let's examine the case of an inter-cluster downlink handoff with the use of a handoff hint. Figure 4.13 (a)–(d) illustrates the procedure for preserving cell sequencing of a VPWA inter-cluster downlink handoff. Note that the numbering of the ATM cell stream in Figure 4.13 is the virtual numbering, which is mainly for helping us to describe how the VPWA handoff scheme can preserve cell sequencing during a handoff. In Figure 4.13 (a), a MH is currently residing in the wireless cell of *BS A.1.3* with a downlink call. The MH moves to a neighboring *BS A.2.1* in Figure 4.13 (b). By extending the original VC route from *BS A.1.3* to *BS A.2.1*, cells of the downlink call can be delivered to the MH by *BS A.2.1*. Suppose the last cell received by the MH before it migrates to *BS A.2.1* is cell 2, cells (starting with cell 3) should be forwarded to *BS A.2.1* and delivered to the MH. This would keep the cells in sequence. A BHO cell (marked as black in Figure 4.13) is sent by *MGS A* after the last cell (cell 7) along the extended *original local route.* After that, *MGS A* delivers all incoming cells (starting with cell 8) of the handoff call to *BS A.2.1* via the *new local route.* As shown in Figure 4.13 (c), cells along the new routing path arriving at *BS A.2.1* before the arrival of the BHO cell are temporarily buffered. When the BHO cell arrives at *BS A.2.1*, the extended *original* *local route* is released and buffered cells are delivered. Figure 4.13 (d) shows the handoff connection is successfully rerouted to *BS A.2.1*, which now serves the MH. Therefore, an end-to-end connection is maintained and so is the cell sequence.

The scheme described above is sufficient in preserving cell sequencing under the assumption that the MH can correctly report the virtual cell number of the last received cell before handoff. Now let's reconsider the case in Figure 4.13. Suppose the last cell received by the MH before handoff is again cell number 2. But the acknowledgment of the receiving of this cell, back to *BS A.1.3* at the WATM data-link sublayer, is lost due to a poor wireless signal during handoff. So, *BS A.1.3* forwards cells to *BS A.2.1*, starting with cell 2. When cell 2 arrives at the MH via *BS A.2.1*, it destroys the cell sequencing of the handoff connection because it introduces duplicate cells in the stream. The duplication is not detectable at the ATM layer. Eventually it causes the entire packet to be dropped at the higher protocol layer, such as SSCOP or TCP.

To overcome the aforementioned problem, we apply the WATM cell format proposed in [81]. Its format is shown in Figure 4.14. WATM cells are used in the WATM sublayer for the wireless communications between a MH and a BS. A WATM cell is slightly modified from an ATM cell. The data payload is still 48-byte long. But the VPI field is removed to make room for an 8-bit cell sequence number. Additionally, a 2-byte cyclic redundancy check (CRC) is added for providing error checking of the entire WATM cell. An acknowledgment cell is used to allow a MH and its serving BS to be synchronized with respect to the sequence number of a call. WATM cells

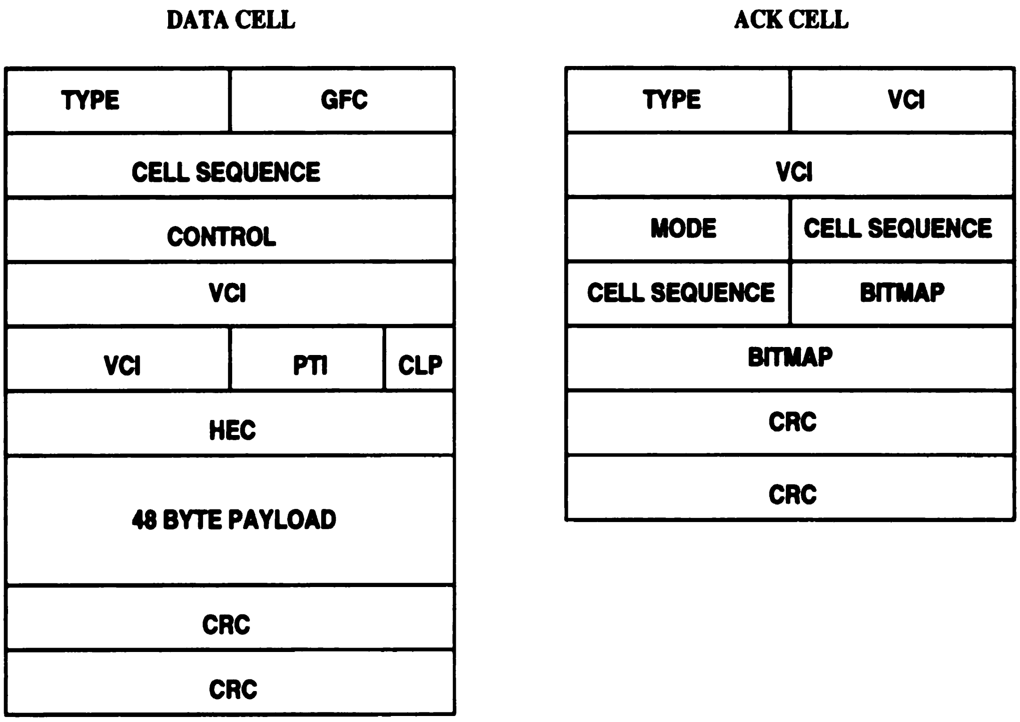


Figure 4.14: Data and ACK formats for WATM sublayer

that move up the ATM layer must be converted back to ATM cells and vice versa. Although cell conversion is required for pass data between the ATM layer and the WATM sublayer, the time delay and the extra cost are both expected to be very small since the conversion process can be done by a simple hardware implementation in an interface card.

Now let's examine the handoff case in Figure 4.13 one more time. Assume that an acknowledgment is issued on a per received cell basis. Again, suppose the last cell received by the MH before handoff in (a) is cell 2. If the acknowledgment for cell 2 is lost, then the acknowledged sequence number on *BS A.1.3* is 1. When the MH sends the signal HO_HINT to *BS A.1.3*, the signal includes the sequence number of the last cell received, which is cell 2. When *BS A.1.3* receives the signal, it updates the sequence number in its system. Subsequent cells are forwarded to *BS A.2.1*, starting with cell number 3. The cell sequence is maintained. If the signal HO_HINT is lost, the signal HO_REQ_SETUP sent by the MH to the NBS (A.2.1) also contains the last received cell sequence number. The sequence number is included in the signal HO_FWD_SETUP which is sent by *BS A.2.1* to *BS A.1.3* to extend the *original local route*. Again, *BS A.1.3* can update the sequence number in its system. Cells with the correct sequence are forwarded to *BS A.2.1* and later delivered to the MH. Once again, the cell sequence of the handoff call is preserved.

4.6 Summary

In this chapter, we presented a VPWA handoff management protocol which is based on the VPWA architecture and an incremental connection re-establishment approach. The signal protocol of the VPWA scheme is based the ATM protocol stack by augmenting several mobility support signals in the signal control layer. A WATM cell format equipped with a cell sequence number is applied in the communications between a MH and a BS at the WATM data-link sublayers. The sequence number of a WATM cell helps to preserve cell sequencing of a handoff connection. According to the structure of the VPWA architecture, connection handoffs are categorized as follow: intra-cluster handoffs; inter-cluster and intra-domain handoffs; and inter-domain handoffs. A signal flow diagram for each handoff case had been illustrated step by step. Based on the buffer requirements and the connection disruption time, analyses were also presented for all cases.

The proposed VPWA handoff mechanism is novel in several ways. First, it can provide fast connection handoffs and reduce call admission control overhead by using the VP technique. Second, the hierarchical network structure of the VPWA architecture eliminates the need for invoking a CCN determining process for intra-domain handoffs. Even in an inter-domain handoff, our VPWA handoff scheme can avoid calling a time-consuming CCN determining function during a handoff with the border MCS design. A CCN determining process is pre-executed before the rerouting of a handoff connection between two domains occurs. Third, the number of lost cells is minimized because a forwarding scheme is available in the VPWA protocol to forward misrouted ATM cells to their destinations. Fourth, our protocol can preserve the cell sequencing of a handoff call. This is made possible by employing an in-band trailing cell, a cell forwarding scheme, sequence numbers of WATM cells, and mobility support signals. Further, with an appropriate network design, there is no additional buffer space needed for connection handoffs at a MGS switch. Finally, the number of signal messages can be reduced by combining signals of the handoff management scheme with signals of the location management scheme.

Chapter 5

Experiments

The main goal of this chapter is to examine our user mobility framework in handling connection handoffs when a mobile host roams in a WATM network. Several programs are developed to implement our VPWA handoff protocol. Experiments are conducted by running the programs on an ATM testbed in conjunction with wireless devices, and the results based on a number of performance metrics are presented.

5.1 Introduction

Due to system limitations and available resources, the network entities (MHs, BSs, and MCSs) in our VPWA infrastructure must be emulated by programs running on workstations or personal computers (PCs).

5.1.1 HSNP ATM Testbed

Our experiments make use of the ATM testbed in the High-Speed Networking and Performance (HSNP) Laboratory in the Computer Science Department of Michigan State University. Figure 5.1 shows the configurations of the HSNP ATM testbed. In the ATM testbed, there are two FORE ASX 200WG ATM switches, one FORE ASX 200BX ATM switch, and two Cisco LightStream 1010 ATM switches. A HP Broadband Series Test System (BSTS) VX-743 is included in the ATM testbed in order to provide network diagnostics and to generate ATM traffic. There are thirteen Sun Ultra Sparc 1s and four high performance PCs (two Pentium II 233 systems and two Pentium Pro 180 dual processor systems) interconnected with each other via multi-mode fiber optics and ATM switches. These workstations and PCs are also connected on a Fast Ethernet network via unshielded twisted pair category 5 (UTP-5) cables and a Cisco Catalyst 2926 Fast Ethernet switch. All Ultra Sparc 1 workstations run under the Sun Solaris 2.5.1 Operating System (OS). A FORE ASX SBA-200E ATM network interface card (NIC) and an ATM device driver are installed on each workstation for connecting to the ATM testbed. Each PC runs under Microsoft Windows NT 4.0 as its default operating system. Efficient Networks ENI-155p-MF ATM network interface cards and device drivers are installed on PCs for connecting them to the ATM testbed. Each PC is also equipped with a 2.4 GHz WaveLAN wireless communications device from Lucent Technologies, Inc. WaveLAN operates in the 2.4220–2.4620 GHz ISM (Industrial, Scientific, and Medical) band.

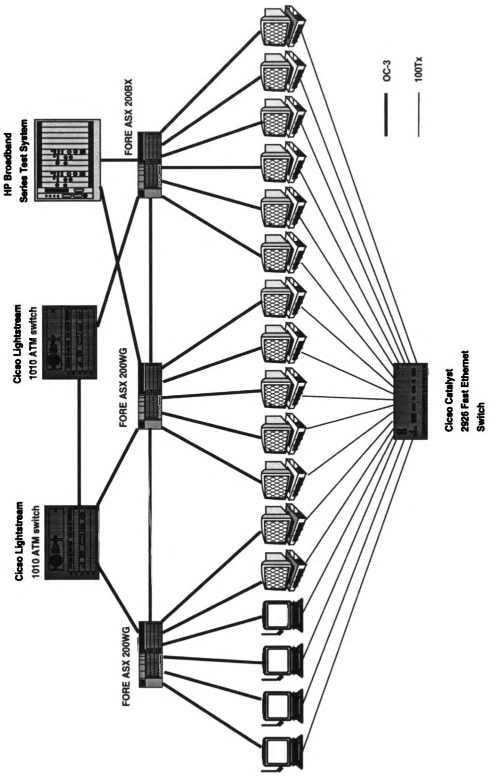


Figure 5.1: Configuration of HSNP ATM testbed

It uses the direct sequence spread spectrum modulation technique to provide reliable communications and protect against eavesdropping.

In our experiments, we need to run Linux OS on the PCs in order to use the *ATM on Linux* package (described later) for controlling ATM NICs. Therefore, these PCs are configured to boot either the Windows NT 4.0 or Linux OS. Linux OS is a free distributed UNIX-clone OS for i386 systems [82]. The Linux system is chosen because (1) it has freely distributable source code, (2) it has quite powerful networking capabilities and (3) some work has already been done for ATM connectivity in this environment.

5.1.2 ATM on Linux Package

ATM on Linux is a software package that contains device drivers for a number of ATM cards, components responsible for ATM signaling, ILMI address registration, IP over ATM [83] and several ATM related protocols [84]. This package was developed by Werner Almesberger from Laboratoire de Reseaux de Communication at the Swiss Federal Institute of Technology in Lausanne, with the help of various contributors on the Internet.

ATM on Linux is composed of ATM protocol stacks and device drivers in the kernel and components in the user space. The user space applications are daemons responsible for ATM signaling and ILMI address registration. Because the protocols used in signaling are rather complex but do most of their work only when setting up or tearing down connections, it is implemented as a daemon in user mode.

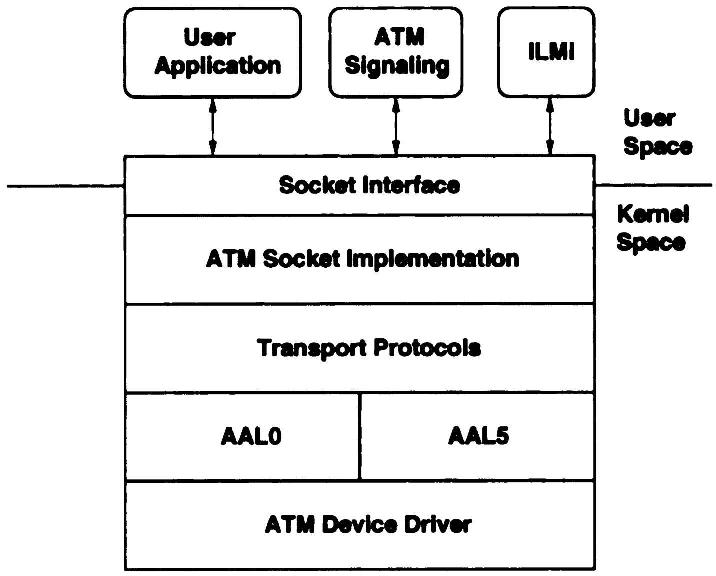


Figure 5.2: ATM on Linux protocol stack

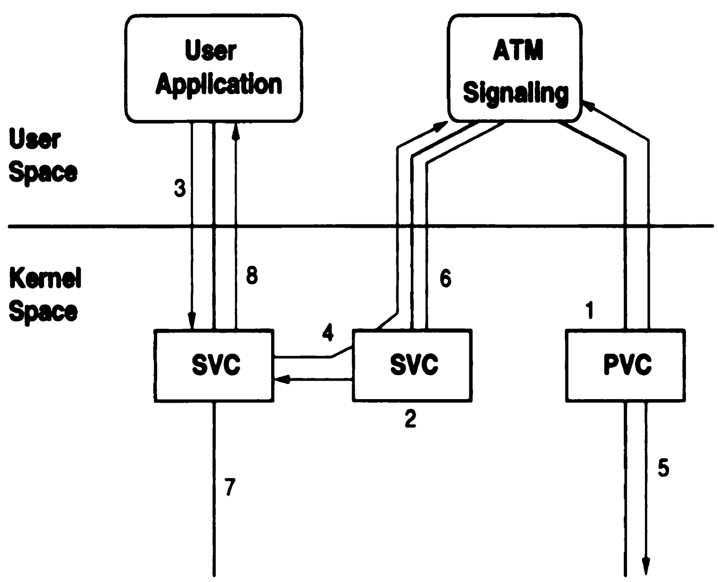


Figure 5.3: ATM on Linux signaling procedure

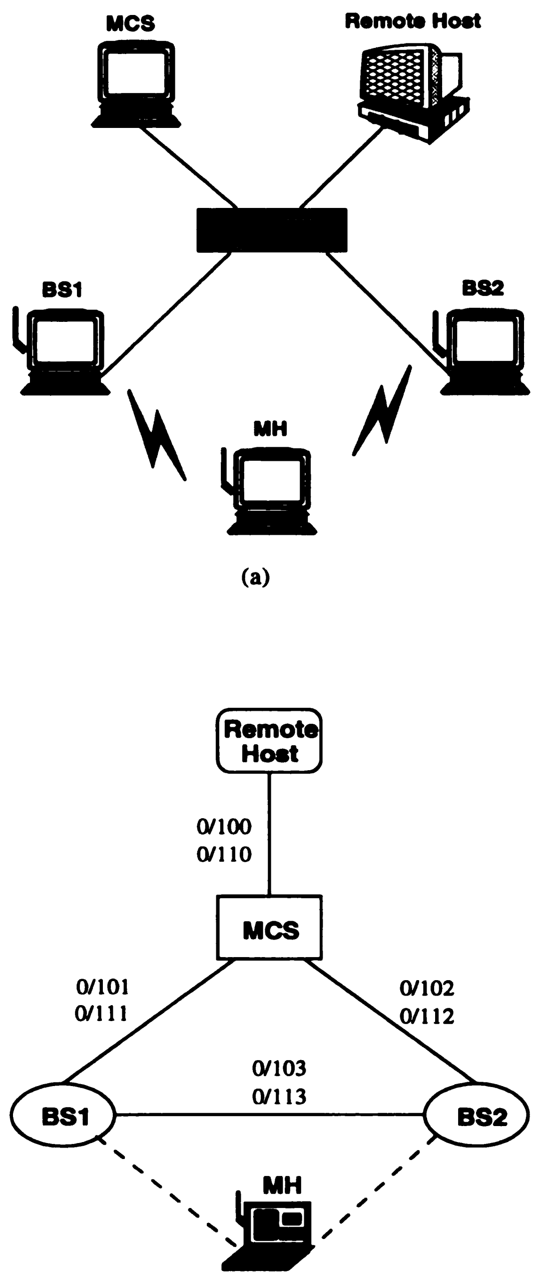
The package provides a set of Application Programming Interfaces (APIs) for accessing the services provided by this software package [85]. The APIs have a format similar to the BSD socket interface. Actually, the APIs are extended from BSD sockets to support ATM interfaces by defining two new socket classes; switched virtual circuit (SVC) and permanent virtual circuit (PVC), to the common socket layer interface. Figure 5.2 shows the protocol stack of the ATM on Linux package. At present, the package only supports AAL0 and AAL5 connections.

Signaling of the ATM on Linux is implemented in the user space as a daemon process. It constructs signaling messages according to requests that are forwarded to it from the kernel and decodes the incoming signaling messages from the signaling channel. The signaling daemon keeps track of open connections and local bindings of service access points. Figure 5.3 illustrates the operations performed by the signaling daemon [86]. When the signaling daemon starts, in (1), it creates a PVC to communicate with the signaling entity in the network via the default signaling channel (VPI=0, VCI=5). A special SVC socket is also created in (2) for exchanging signaling messages with the kernel. When a user application requests a connection to a remote ATM end system in (3), the signaling daemon is notified of the request by the kernel in (4). After receiving the request, the signaling daemon performs the standard connection setup process (follows either UNI 3.0 or 3.1) with the ATM network in (5). Once the connection is established, the signaling daemon informs the kernel of the availability of the connection in (6). The kernel then sets up the local part of the data connection in (7) and notifies the application in (8). A similar procedure is performed for an incoming call.

5.2 Experimental Setup

In the experiments, we investigate intra-cluster connection handoffs. Figure 5.4 (a) displays the experimental setup on the HSNP ATM testbed while Figure 5.4 (b) presents the virtual topology of the experimental setup. In the experiments, a PC is functioning as a mobile host. Two other PCs perform the tasks of base stations. The MH communicates with these two BSs via WaveLAN wireless devices. The BSs connect to the same MCS (another PC) via ENI-155p-MF ATM NICs, multiple mode fiber optics, and a FORE ATM switch. Communications between BSs and the MCS is via the ATM network. All PCs run the Linux OS in conjunction with the ATM on Linux package. A Sun Ultra Sparc 1 is used as a remote host on the ATM backbone. As it can be seen in Figure 5.4 (a), the remote host is physically connected to the same ATM switch to which all aforementioned PCs are connected. Data communications are performed between the MH and the remote host through all the experiments.

Figure 5.4 (b) shows that two PVCs are pre-established on each link between two communication nodes. One of the PVCs is for signaling because we have no control on the default signaling channel. The other PVC is for data communications in the experiments. Details of using these PVCs in conducting the handoff experiments is presented later in this chapter.



(b)

Figure 5.4: Experiment setup: (a) the physical topology (b) the virtual topology

5.2.1 Architecture of VPWA Network Entities

Since the source code associated with the equipments on the ATM testbed is not available, we need to develop software for performing the functionalities of all the network entities of the experiments. These network entity programs include the desired VPWA handoff protocol to support user mobility. The architecture of these network entities is shown in Figure 5.5. At a MH, the MSU AAL layer is a simplified standard ATM AAL layer. Its main task is to perform the segmentation and assembly of data packets exchanged between the user application and the MSU ATM layer. Most of the ATM functions are implemented in the MSU ATM layer. Segmented data from the MSU AAL layer are encapsulated with ATM cell headers. The MSU ATM layer handles data coming from the lower MSU WATM sublayer and then passes data to the upper MSU AAL layer. The MSU WATM sublayer is used to perform communications between wireless links. An ATM cell from the MSU ATM layer is converted into the format of a WATM cell as shown in Figure 4.14 of Chapter 4. This layer also applies a simple reliable communication scheme, which requires the receiver to send acknowledgments back to the sender. At a base station, the protocol stack supports both a WATM interface and an ATM interface. Signaling messages have to traverse up to the MSU signaling layer while data cells are handled at the ATM layer. The ATM layer at a BS transports data cells between a WATM interface and an ATM interface. Similarly, the MSU signaling layer processes and forwards signals between these two interfaces. Note that, at the MCS and the remote host, no WATM sublayer is needed because they only have ATM interfaces.

In the development of these network entity programs, the previously mentioned layering features, including our VPWA handoff protocol, are implemented as a set of functions, called Mobility Support Utilities (MSU). The MSU functions are compiled and prepared as a library, called **msu_lib**, to be linked by these programs during compilation. The library provides generic functions to open and close a connection, on either a WATM or an ATM interface. Most signaling functions are also included in this library. Figure 5.6 illustrates the process structure of these network entity programs.

At a MH, a user application is a process which communicates with the signaling daemon **mh_sigd** via a UNIX domain UDP socket. A user application process instructs the signaling daemon to open and close a connection via this socket. All signaling operations are executed at the **mh_sigd** daemon. The daemon presents a virtual WATM interface to the user application by hiding the underline WaveLAN wireless device from the application layer. This is made possible by maintaining a VC table, which maps a VC at the MSU ATM layer to a UDP socket at the lower layer.

In default, a socket is setup for signaling between the signaling daemon **mh_sigd** at a MH and the signaling daemon **bs_sigd** at a BS. This signaling socket is used as the default signaling channel between two ATM interfaces, according to the ATM standard. At a BS, two daemons, **bs_sigd** and **atmsigd**, run cooperatively to perform the duty of a BS. The former is the main signaling daemon, which handles all the signaling messages, manipulates connections on the WATM interface, and maintains

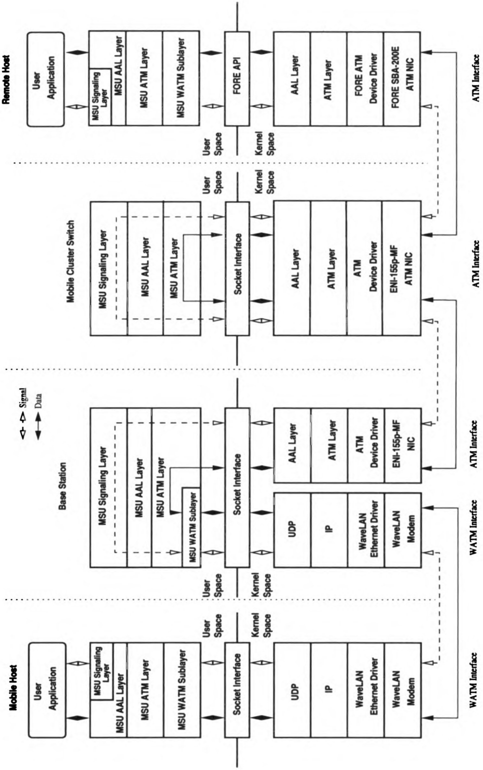


Figure 5.5: Architecture of the VPWA protocol stack with the MSU functions

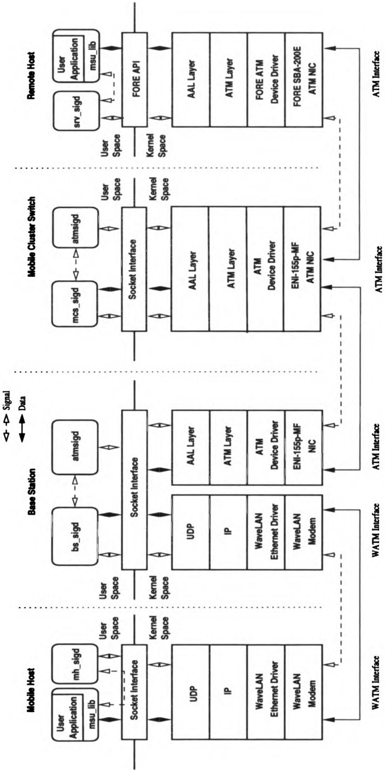


Figure 5.6: Process structure of the network entity programs

a VC table. Similar to the VC table at a MH, the VC table at a BS maps sockets in the kernel to VCs on both the WATM and the ATM interfaces. It also links the VC on the WATM interface to a VC on the ATM interface. Therefore, a virtual connection is established between these two interfaces. The **atmsigd** daemon, on the other hand, is the signaling daemon from the ATM on Linux package. In the experiments, this daemon is essentially used to open and close a real ATM VC on the ATM interface. The **bs_sigd** instructs the **atmsigd** to manipulate connections on the ATM interface by using the appropriate API of the ATM on Linux package.

Similar to the BS, the MCS has two daemons (**mcs_sigd** and **atmsigd**) working together. Again, the **mcs_sigd** daemon is the main process of handling signaling messages. The **atmsigd** is merely used to perform the operations of opening and closing ATM VCs according to the instructions from **mcs_sigd**. Unlike the **bs_sigd**, no WATM sublayers are included in the **mcs_sigd**. Note that, in Figure 5.4 (b), the MCS connects to two BSs and a remote host. It requires three ATM interfaces to connect to these three network entities. However, in the real setup shown in Figure 5.4 (a), the MCS has only one physical ATM interface. To resolve this problems, **mcs_sigd** provides multiple virtual ATM interfaces by maintaining a data structure for each virtual interface. Each virtual interface has its own ATM VC table. A connection across different interfaces is constructed by linking virtual channel entries from different virtual interfaces together.

At the remote host, as it can be seen in Figure 5.5, the protocol structure is the same as that of the MH. The user application at the remote host is also similar to the user application at the MH. A user application relies on the signaling daemon **srv_sigd** to perform the desired signaling operations. In the physical experimental setup, the remote host is a Sun Ultra 1, equipped with a FORE SBA-200E ATM NIC. The **srv_sigd** daemon applies the API provided by FORE Systems to instruct the ATM device driver in the kernel to open and close connections on the ATM interface.

5.2.2 Call Setup

Now let's consider an uplink call initiated by the MH to the remote host. Before a user application at a MH can communicate with its communication party on the remote host, an end-to-end connection must first be setup. A user application issues a connection setup request by calling a **msu_open()** function in the **msu_lib**. This **msu_open()** function is a script wrapper for opening a generic connection, which can be either a WATM connection or an ATM connection. Here, at the MH, this function opens a UDP connection via the socket interface. The kernel prepares a UDP socket and returns a socket descriptor back to the user application process. When **msu_open()** is returned from the kernel successfully, a function **msu_connect()** is called. The **msu_connect()** function sends a connection request to the **mh_sigd** daemon. Upon receiving the request, the daemon generates a modified **SETUP** signal, which is based on the ATM UNI-3.1 signaling standard. This signal message is sent to the kernel via the signaling socket. The signal message descends down the protocol layers in the kernel space and reaches the WaveLAN modem, and then it is delivered to the serving BS through the air.

At the serving BS, the signal traverses upward from the physical layer (WaveLAN modem) to the UDP layer in the kernel. Later, the signal is retrieved by the signaling daemon **bg_sigd** in the user space by reading the signaling socket. As shown in Figure 5.5, the signal moves up the protocol layer in the user space until it reaches the MSU signaling layer. Once it reaches this layer, a CAC process should be invoked to determine whether to accept the request or not. In our experiments, the request is assumed to be a UBR traffic which should always be accepted. Hence, we control the demanded bandwidth in the experiments not to exceed the link bandwidth. Therefore, no CAC function is needed. Before the connection request signal is forwarded to the MCS, a connection data structure, shown in Figure 5.7 is created to handle the connection at the BS. The connection control block stores the essential information about a connection. For example, the call reference number is used to identify the call when **bs_sigd** receives signaling messages from other network entities. The BS forwards the call setup signal by sending it down the protocol layers. Again, the signal traverses down the AAL layer to the physical layer (the ENI-155p-MF ATM interface card), and later sent to the MCS via fiber optics and an ATM switch.

When the setup request arrives at the signaling daemon **mcs_sigd** of the MCS, the MCS makes a routing decision and creates a connection data structure for this connection. It then forwards the request to the remote host. When the signaling daemon **srv_sigd** at the remote host receives the request, it informs the target application via sending a local signal. After the user application accepts the request, the signaling daemon sends a confirmation signal (**CONNECT**) back to the MH.

```
struct msu_vcb{
   call_state        state;     /* call state */
   atm_vcb           a_vcb;     /* control block on ATM */
   watm_vcb          w_vcb;     /* control block on WATM */
   msu_nsap          scrn;      /* source NSAP */
   msu_nsap          dstn;      /* destination NSAP */
   struct msu_qos    qos;       /* QoS */
   int               call_ref;  /* call reference no. */
   bs_info           nbs;       /* new BS info */
   bs_info           cbs;       /* current BS info */
   int               sflag;     /* signal flag */
   int               direction; /* flow direction */
   msu_itf           *in_itf;   /* input interface */
   msu_itf           *out_itf;  /* output interface */
   msu_vcb           *prev;     /* pointer to previous vcb block */
   msu_vcb           *next;     /* pointer to next vcb block */
};
```

Figure 5.7: Data structure of a connection control block at a BS

The confirmation signal traverses along the signaling channel in the opposite direction. Upon receiving the confirmation signal, **mcs_sigd** checks its VC table and binds a pre-setup PVC data channel for the connection on each virtual interfaces. The signal is then transported to the BS. The same process is incurred at the BS before the signal is sent to the MH. At the MH, after receiving the **CONNECT** signal, **mh_sigd** informs the user application of the completion of the connection setup. Thus, the connection setup is completed and the data communications between two user applications can proceed.

5.2.3 Data Delivery

During the data communications phase, the user application at the MH generates data packets and hands them to the MSU AAL layer. These data packets are segmented into 48-byte pieces. According to the routing information of the connection, a 5-byte

ATM cell header is concatenated with the 48-byte payload into an ATM cell. These ATM cells are delivered to the WATM sublayer where the 48-byte payload is kept intact but the header is replaced with a 4-byte header. A 2-byte CRC trailer is added after the data payload for error checking on the entire WATM cell. As shown in Figure 4.14 of Chapter 4, each WATM cell is equipped with a unique cell sequence number in a connection. The sequence number of a connection must synchronize with that one at the serving BS.

In the real world, a WATM cell is encapsulated in several protocol packets in the kernel before it is sent across the air. Every encapsulation adds an additional protocol header to a WATM cell. Generally, a WATM cell has at least 28 bytes of additional headers, including 8 bytes of UDP header and 20 bytes of IP header [87]. At the WaveLAN Ethernet driver, several WATM cells are put in a single Ethernet frame (which has a 14-byte frame header and a 1500-byte maximum frame payload) before being sent to the serving BS.

At the BS, the data is recovered into WATM cells in the kernel before being delivered to the MSU WATM sublayer in the user space. Acknowledgments are sent back to the MH to indicate the successful receptions. These WATM cells are converted back to ATM cells before they enter the MSU ATM layer. Similar to the VPI/VCI table lookup on an ATM switch, a table lookup is performed at the BS. Based on the information at the lookup table, these ATM cells are pipelined to the VC on the ATM interface. In our implementation, the data structure of the connection created at the connection setup phase stores the UDP socket descriptor and the VPI/VCI on the WATM side. The entry is linked to its partner entry (which has the ATM socket descriptor and the VPI/VCI information) on the ATM interface. Based on the information of the data structure, these ATM cells are pushed back into the kernel space through the ATM data socket. These ATM cells are then sent to the MCS via the ATM network.

A similar process is performed when these ATM cells arrive at the MCS. ATM cells are delivered to the remote host (the destination) by the MCS. When these ATM cells arrive at the remote host, they are retrieved by the **srv_sigd** daemon from the kernel. Later, they are converted back to the format of user data packets before being delivered to the target user application.

During the transition of a connection handoff, multiple input data structures of a connection temporarily co-exist; although, only one is active. For example, an uplink handoff connection has a data structure on each interface to the NBS and the CBS. The data structure on the NBS remains inactive until the MCS detects the BHO cell and removes the data structure on the CBS side. During this period of time, cells from the side of NBS are buffered in the input queue of the data structure.

Note that all ATM connections for data communications in the experiments are setup as AAL0 connections. An AAL0 connection allows us to bypass the processing overhead in the AAL layer and transmit ATM cells directly. Therefore, in-band BHO cells in the VPWA protocol can be sent and received by the network entity programs. In the transmission of an ATM cell, only the first 4 bytes of the cell header and the 48-byte payload are needed for calling the transmission function provided by the ATM on Linux or the FORE API. The fifth byte of the cell header (HEC) is provided by hardware on the ATM NIC. On the other hand, signaling channels in our experiments are setup as AAL5 connections. This matches the requirements for the ATM signaling standard.

5.3 Experimental Results

Experiments were performed according to the experimental setup described in the previous section. During the experiments, the handoff events of the MH were generated by a handoff module in the program of the MH network entity. The MH moved between the two BSs shown in Figure 5.4 (b). The experimental data we collected was the connection disruption time and the buffer requirements at both the CBS and the NBS. The connection disruption time is a very important performance metric since most users evaluate the quality of a PCS service based on the communication quality. The buffer requirements on the system were the index to estimate the cost of implementing such a system. The cell sequencing at each handoff is also checked in all experiments.

Table 5.1 shows the parameters and measurements of sample downlink experiments. The number of total lost data packets shows that if the cell forwarding of our protocol is not used, this number of data packets will be dropped and retransmitted. With the cell forwarding mechanism, no cell was trapped at base stations through all experiment runs. Thus, no data packet was dropped and the cell sequencing was

Parameters	set 1	set 2	set 3	set 4
Packet size (bytes)	256	512	1024	2048
Total data packets sent	1500	1500	1500	750
Measurement				
Total experiment time (sec)	150.299	150.349	150.443	135.443
Total handoff events	17	17	17	14
Total data packets dropped	4	8	9	14
Mean connection disruption time (msec)	143.155	106.334	138.276	142.651
Mean buffered length at CBS (cells)	0.411	0.941	2.35	2.5
Mean buffered length at NBS (cells)	9.41	13.3	31.9	42.6
Max buffer length at NBS (cells)	24	32	84	105
Mean throughput (Kbps)	25.406	46.562	93.067	101.109
Mean handoff interval (sec)	9.29	9.287	9.288	9.47

Table 5.1: Experimental results in various parameter sets

preserved. Therefore, unnecessary retransmissions can be avoided and the throughput can be increased. Also shown in Table 5.1, the total dropped packets are highly related to the handoff frequency and the traffic load. With data rate about 101 Kbps, there is at least one packet dropped for each connection handoff. The mean cell buffer size increases as the traffic load and the connection disruption time increases. The average connection disruption time (time delay to complete a handoff session) is around 105–145 milli-seconds.

Figures 5.8–5.11 depict traces of ATM cell stream from the remote host to the MH based on different packet sizes and traffic loads. Most of the time, cells arrive steadily at the MH. The spikes on these figures are due to buffered cells at the NBS being flushed to the MH, once a connection handoff is completed. From Table 5.1, the mean buffer space needed at the NBS and the CBS is 2.5 cells and 42.6 cells, respectively, when the traffic is around 101 Kbps. Even when a maximum observed value of 105 cell

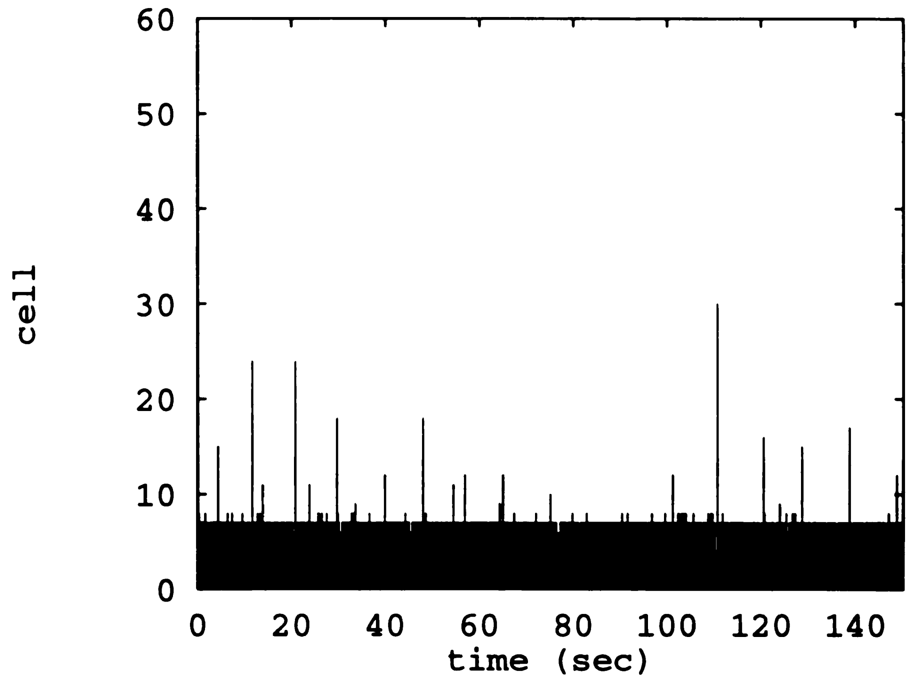


Figure 5.8: Trace of ATM cell stream with parameters of set 1 (every 100 msec)

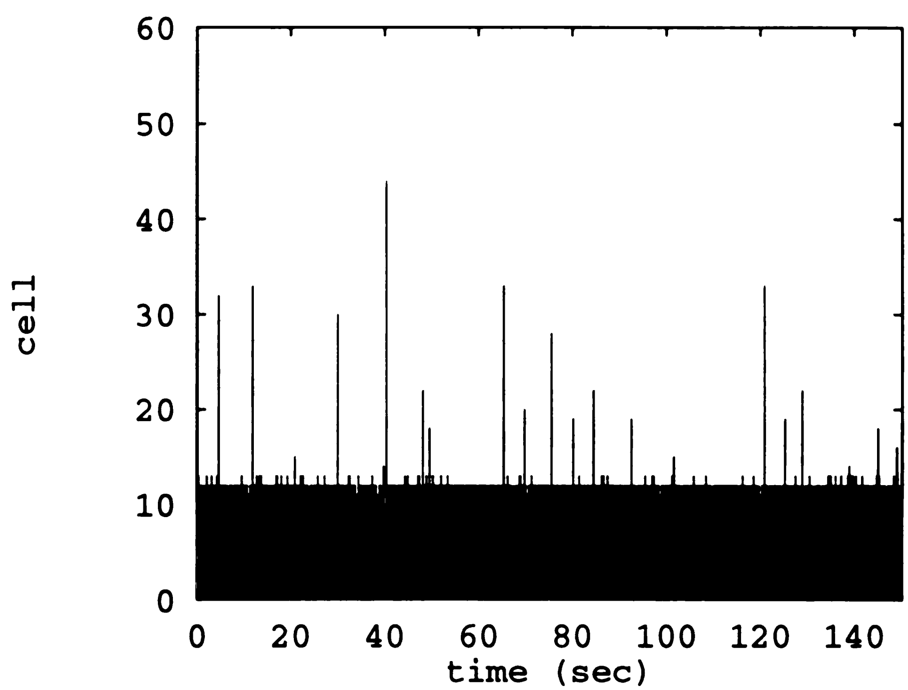


Figure 5.9: Trace of ATM cell stream with parameters of set 2 (every 100 msec)

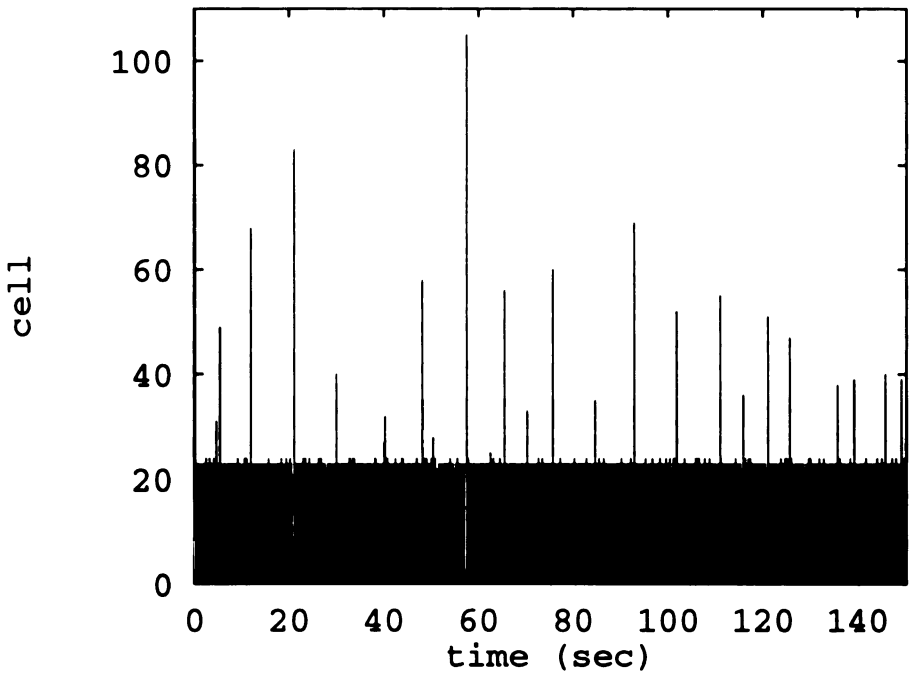


Figure 5.10: Trace of ATM cell stream with parameters of set 3 (every 100 msec)

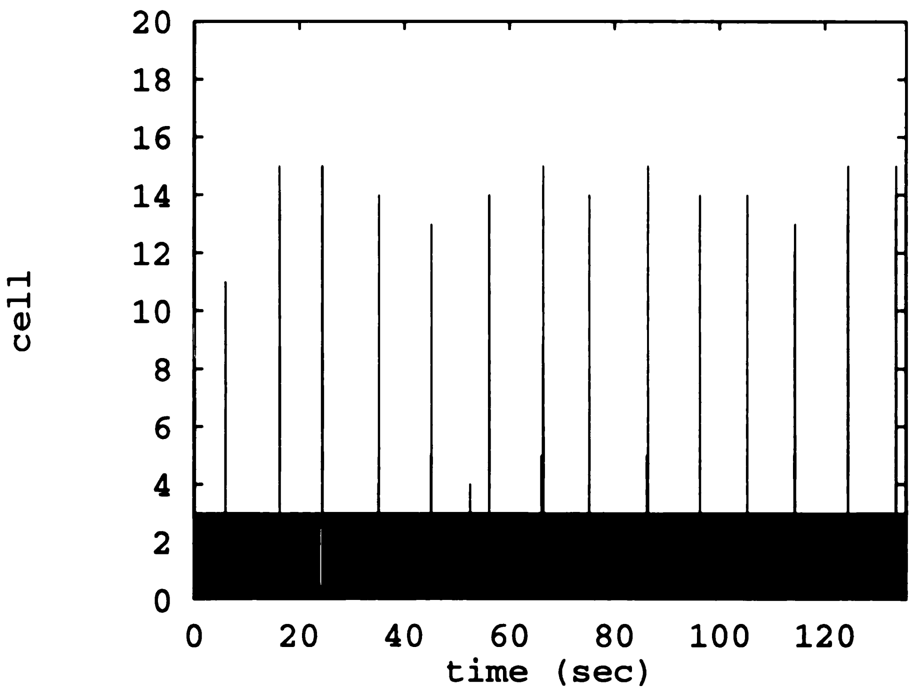


Figure 5.11: Trace of ATM cell stream with parameters of set 4 (every 10 msec)

buffers is needed at the NBS, this is considered a small price to pay since this scheme greatly reduces the buffer requirements at both communication ends, with respect to packet retransmissions. Hence, the buffer requirements can be further reduced when the connection disruption time is reduced. Through all experiments, the cell sequence of a handoff connection is preserved with the use of our VPWA handoff protocol. In turn, it reduces the volume of retransmission traffic.

The experiments conducted here are mainly to examine our VPWA protocol. The experimental results should only be considered as reference values when examining the performance of the VPWA mechanism for the future WATM network. The connection disruption time and the buffer requirements can be substantially reduced by enhancing both hardware and software. For example, at a BS, delivering an ATM cell from the ATM interface to the WaveLAN wireless interface involves four copy operations in our experiments. The first copy is from the ATM interface card to the kernel space. The second copy is from kernel to the signal daemon in the user space. The ATM cell is then converted to a WATM cell. Then the third copy moves the WATM cell from the user space in the kernel again. Finally, the WATM cell is copied to the WaveLAN interface card. All these copy operations require CPU interventions. Hence, the two copy operations between the user space and the kernel space involve the context switching overhead because the Linux is a general purpose OS. Also, the protocol processing (UDP layer, IP layer, and Ethernet layer) in the kernel contributed a tremendous amount of overhead to the experiments. Therefore, with appropriate software and hardware support, the connection disruption time and the buffer requirements should be significantly reduced.

5.4 Summary

In this chapter, we have described the experiments for testing our VPWA handoff protocol in supporting connection handoffs. First, we presented the experimental setup on the HSNP ATM testbed. Four PCs, an Ultra Sparc 1 workstation, and a FORE ATM switch form the physical experiment topology. Several programs were developed to function as the network entities in our VPWA architecture. These programs implement the augmented mobility support functions in the VPWA handoff protocol.

The PCs in the experimental setup run the Linux OS in order to make use of the ATM on Linux package. The ATM on Linux is a free available software which contains a device driver for a number of ATM NICs and standard ATM protocol components. Our network entity programs use this package to physically open or close virtual connections on the ATM interface card. The VCs for data connections in the experiments are the AAL0 type of connection. This allows us to send and receive ATM cells directly. So in-band BHO cells in the VPWA protocol can be detected by the programs.

We have conducted several experiments and collected the experimental results. Throughout all experiments, connection handoffs are handled smoothly and the cell sequencing is preserved. The experimental results also show that if our VPWA protocol is not used, data packets are dropped during connection handoffs even if the connection carries low bandwidth traffic. A dropped data packet causes the QoS degradation for time-sensitive traffic or requires retransmission for loss-sensitive traffic. Although there is some overhead (buffer space) involved in the VPWA approach, it is considered advantageous since the VPWA approach can avoid the QoS degradation on time-sensitive traffic and simultaneously minimize buffer requirements at end systems for loss-sensitive traffic.

Chapter 6

Summary and Future Work

In this chapter, we summarize the major contribution of the research in this dissertation and outline some directions for future work.

6.1 Summary

The parallel trends of demanding multimedia-capable PCS services have sparked integration of voice, video, image, and data in both telecommunication and computing environments. To effectively integrate wireless and wireline networks, wireless ATM networks are considered to be one of the best scenario. Among issues related to WATM networks, handling user mobility is considered the most challenging one. In this thesis, we introduce a framework of handling user mobility in WATM networks. This framework has three essential components: a VPWA network architecture, a VPWA location management scheme, and a VPWA handoff management mechanism.

The VPWA network architecture separates the wireless peripheral network from the ATM backbone network. So, most mobility support functions are implemented on the peripheral network. By doing this, major modifications on the current ATM network can be avoided when deploying the WATM network. Also, the use of the virtual path technique to construct the peripheral infrastructure allows network managers to configure the network topology according to the changes of the network traffic.

The VPWA location management scheme is designed to complement the current ATM routing scheme by processing queries for locations of mobile hosts and delivering calls to mobile hosts on WATM networks. The VPWA location management scheme has the following advantages:

- the inefficient location management problem in the IS-41 model is avoided.
- it works cooperatively with the current ATM routing mechanism for setting up connections.
- it helps to migrate the PCS services from the current PCS network to the ATM based network smoothly.

The VPWA handoff management scheme is designed to maintain the continuity of a call at a mobile host when the mobile host is roaming in the infrastructure of the WATM network. This handoff scheme can also handle a high frequency of connection handoffs in WATM networks. The major advantages of the VPWA handoff management scheme are listed as follows:

- it provides fast connection handoffs and reduces call admission control overhead.

- the need for invoking a CCN determining process for intra-domain handoffs is eliminated.

- calling the CCN determining process during an inter-domain handoff is avoided by using the border MCS.

- a cell forwarding mechanism in the VPWA handoff scheme avoids misrouted cells from being discarded during a connection handoff.

- the cell sequencing of a handoff call is preserved by the proposed scheme. This helps to avoid the QoS degradation of time-sensitive traffic and eliminate excessive data retransmissions of loss-sensitive traffic.

- the number of signal exchanges is further reduced by combining the signals of the handoff scheme with the signals of the location management scheme.

We have conducted several experiments based on our VPWA user mobility framework over an ATM testbed. The experimental results provide reference performance values when implementing such systems. In summary, we consider the followings as our major contributions:

- a network architecture is introduced to avoid major modifications on current deployed ATM equipment when constructing the future WATM network.

- mobility support functions are designed to fit in the ATM protocol without violating the ATM standard.

- an efficient location management scheme is proposed to complement the ATM routing mechanism for tracking mobile user locations and setting up calls.

- a fast handoff mechanism is presented to reroute handoff connections and to preserve their cell sequencing.

6.2 Future Work

In this thesis, we provide a framework for handling user mobility on WATM networks. However, there are still issues that need to be addressed. Some possible directions for future work to further enhance this research are as follow:

- **Improving signaling processing delay:** If our augmented signaling functions can be inserted in the ATM protocol stack and implemented in real systems, they should dramatically improve the signaling processing capability. Moreover, to further improve the signaling processing time, a realtime OS should be used to effectively handle signals. If this were done, the signaling copy operations and context switchings in our experiments could be eliminated. Additionally, the connection disruption time and the buffer space could be reduced substantially.

- **Minimizing data delivery delay on the critical path:** In our experiments, there are four copy operations involved at each BS and each MCS to deliver a cell from one interface to another interface. These operations incur an excessive amount of overhead. The delay of transporting cells at a BS should be greatly reduced by applying a special hardware and a special memory mapping technique. Special hardware can be used to quickly convert cells between the WATM sublayer and the ATM layer. The memory mapping technique is able to map data cells on the memory banks of two interface cards in the kernel space, thus eliminating the need for the copy operation between the kernel and the interface card. Hence, by using a scatter-and-gather send function, the copy operation between the two interface cards can be eliminated. Ideally, the copy operation can be reduced to one and the buffer space in the system can be minimized.

- **Exploring the WATM Sublayer:** The WATM sublayer handles the wireless communication segment of a WATM connection. Since the wireless segment is the performance critical point of a WATM connection, the design of the WATM sublayer is the critical. Different wireless air interfaces (TDMA or CDMA) have various advantages and constraints. It is an important topic to investigate the tradeoffs for the wireless data-link sublayer and the wireless MAC sublayer under different physical wireless technologies. Further, the performance study on these various wireless technologies in conjunction with our VPWA user mobility framework provides another interesting research direction.

- **Investigating TCP performance over WATM networks:** Because many networking applications today are based on the TCP/IP protocols, it is inevitably necessary to support the TCP protocol over WATM networks. The study of TCP performance over a WATM network is highly desirable. The TCP performance over a wireless network is expected to be very poor since lost packets are assumed to be a result of congestion and a *slow start* is turned on at the sender [88, 89]. The problem should become worse as the handoff frequency is expected to become higher in the future WATM networks. Applying our VPWA handoff management protocol should relieve the TCP performance problem. However, the improvement of applying the VPWA scheme needs further study. Also, it may be an interesting topic to compare the performance of a mobile-IP approach with the performance of the VPWA approach.

- **Studying QoS guarantees for multimedia applications over WATM networks:** Providing QoS guarantees for multimedia applications on the WATM network is another challenging issue that needs to be addressed. The WATM QoS issue is more complex than the QoS issue on ATM networks, because the QoS contract may be violated when mobile users are roaming in WATM networks. Connections with violated QoS will be either forced to terminate or degrade. Hence, the forced termination rate is highly related to the new call blocking rate. The tradeoff between guaranteeing the QoS of on-going calls and accepting new calls is important in the study of providing the QoS guarantee for PCS services on WATM networks.

APPENDICES

Appendix A

Glossary

AAL	ATM Adaptation Layer
ABR	Available Bit Rate
AMPS	Advanced Mobile Phone System
ANSI	American National Standards Institute
API	Application Programming Interface
ARQ	Automatic Repeat reQuest
ATM	Asynchronous Transfer Mode
BHO	Before Handoff
B-ISDN	Broadband-Integrated Services Digital Network
BS	Base Station
BSC	Base Station Controller
BSTS	Broadband Series Test System
CAC	Call Admission Control
CATV	Cable Television
CBR	Constant Bit Rate
CBS	Current Base Station
CCN	Common Cross Node
CDMA	Code Division Multiple Access
CDVT	Cell Delay Various Time
CRC	Cyclic Redundancy Check
NIC	Network Interface Card
DCS	Digital Cellular Service
EIA/TIA	Electronic/Telecommunications Industry Associations
E-mail	Electronic Mail
ETSI	European Telecommunications Standards Institute

FAWN	Flexible Adaptor for Wireless Networking
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FIFO	First Come First Out
FPLMTS	Future Public Land Mobile Telecommunications System
FTP	File Transfer Protocol
Gbps	Giga Bit Per Second
GHz	Giga Hertz
GSM	Global System for Mobile Communications
HEC	Header Error Control
HLR	Home Location Register
HSNP	High-Speed Networking and Performance
ILMI	Interim Local Management Interface
IP	Internet Protocol
IR	Infrared
IS-41	Interim Standard 41
ISM	Industrial, Scientific, and Medical
ITU	International Telecommunications Union
Kbps	Kilo Bit Per Second
LAN	Local Area Network
MAC	Media Access Control
Mbps	Mega Bit Per Second
MCR	Mean Cell Rate
MCS	Mobile Cluster Switch
MDR	Multiservice Dynamic Reservation
MGS	Mobile Gateway Switch
MH	Mobile Host
MPEG	Motion Picture Expert Group
MSC	Mobile Switching Center
MSU	Mobility Support Utility
NBS	New Base Station
NCNR	Nearest Common Node Rerouting
NIC	Network Interface Card
NSAP	Network Service Access Point
OAM	Operation and Maintenance
OS	Operating System
PACS	Personal Access Communication System
PAN	Permanent ATM NSAP
PBS	Portable Base Station
PC	Personal Computer
PCN	Personal Communications Network
PCS	Personal Communications Service
PCR	Peak Cell Rate

PNNI	Private Network to Network Interface
PSTN	Public Switching Telephone Network
PVC	Permanent Virtual Channel (Circuit)
QoS	Quality of Service
RATM	Radio ATM
RM	Resource Management
RF	Radio Frequency
SCR	Sustained Cell Rate
SONET	Synchronous Optical Network
SS7	Signaling System No. 7
SSCOP	Service Specific Connection Oriented Protocol
STP	Signal Transfer Point
SVC	Switched Virtual Channel (Circuit)
SWAN	Seamless Wireless ATM Network
TAN	Temporary ATM NSAP
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TLDN	Temporary Local Directory Number
UBR	Unspecified Bit Rate
UDP	User Datagram Protocol
UHF	Ultra High Frequency
UID	User Identification
UMTS	Universal Mobile Telecommunications System
UNI	User-Network Interface
UTP-5	Unshielded Twisted Pair Category 5
VBR	Variable Bit Rate
VBR-RT	VBR Realtime
VBR-NRT	VBR Non-Realtime
VC	Virtual Channel(Circuit)
VCI	Virtual Channel (Circuit) Identifier
VP	Virtual Path
VPI	Virtual Path Identifier
VPWA	Virtual Path WATM Architecture
VLR	Visitor Location Register
VLSI	Very Large Scale Integration
WATM	Wireless Asynchronous Transfer Mode
WWW	World Wide Web

Appendix B

Performance Derivation

Parameters and notations used in the model analysis for deriving the required performance are defined in Table 4.2 of Chapter 4.

B.1 Intra-Cluster Uplink With-Hint Analysis

T_1	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_1	=	T_1
T_2	=	$(TD_c + PD + PPT_x)N_h$	RT_2	=	$RT_1 + T_2$
T_3	=	$(TD_c + PD + PPT_x)N_a$	RT_3	=	$RT_1 + T_3$
T_4	=	$T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	RT_4	=	T_4
T_5	=	$(TD_c + PD + PPT_x)N_b + PPT_{cac}$	RT_5	=	$min\{RT_2, RT_4\} + T_5$
T_6	=	$(TD_c + PD + PPT_x)N_b$	RT_6	=	$RT_5 + T_6$
T_7	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_7	=	$RT_6 + T_7$

Connection disruption time at MH:

$$
\begin{aligned}
Tdisrupt &= RT_7 \\
&= T_7 + T_6 + T_5 + min\{T_2 + T_1, T_4\} \\
&= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + PPT_{cac} + 2N_b(TD_c + PD + PPT_x) + min\{N_h(TD_c + PD
\end{aligned}
$$

$+PPT_x), (T_{acq} + PPT_{cac})\}$

Buffer requirement at MCS:

$$RT_{BHO} = RT_1 + (TD_d + PD)N_a + PPT_x$$

$$RT_{n1} = RT_7 + (2PPT_{cwa} + TD_{wlc} + PD_{wl}) + N_b(TD_d + PD)$$

$$\begin{aligned} B_{CCN} &= max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\}BW_{vci} \\ &= max\{[RT_1 + (TD_d + PD)N_a + PPT_x] - [RT_7 + (2PPT_{cwa} + TD_{wlc} + PD_{wl}) + N_b(TD_d + PD)] \\ &\quad + T_{sw}, 0\}BW_{vci} \\ &= max\{[T_{sw} + (N_a - N_b)(TD_d + PD)] - [4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + PPT_{cac} + 2N_b(TD_c + PD \\ &\quad + PPT_x) + min\{N_h(TD_c + PD + PPT_x), (T_{acq} + PPT_{cac})\}], 0\}BW_{vci} \end{aligned}$$

In general, $RT_{BHO} < RT_{n1}$, so $B_{CCN} = 0$.

B.2 Intra-Cluster Uplink Without-Hint Analysis

$T_1 = T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	$RT_1 = T_1$
$T_2 = (TD_c + PD + PPT_x)N_h$	$RT_2 = RT_1 + T_2$
$T_3 = (TD_c + PD + PPT_x)N_a$	$RT_3 = RT_2 + T_3$
$T_4 = (TD_c + PD + PPT_x)N_b + PPT_{cac}$	$RT_4 = RT_1 + T_4$
$T_5 = (TD_c + PD + PPT_x)N_b$	$RT_5 = RT_4 + T_5$
$T_6 = PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	$RT_6 = RT_5 + T_6$

Connection disruption time at MH:

$$\begin{aligned} Tdisrupt &= RT_6 \\ &= T_6 + T_5 + T_4 + T_1 \\ &= T_{acq} + 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 2PPT_{cac} + 2N_b(TD_c + PD + PPT_x) \end{aligned}$$

Buffer requirement at CCN:

$$\begin{aligned}
RT_{BHO} &= RT_2 + (TD_d + PD)N_a + PPT_x \\
RT_{n1} &= RT_6 + (PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa}) + (TD_d + PD)N_b \\
B_{CCN} &= max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\}BW_{vci} \\
&= max\{[RT_2 + (TD_d + PD)N_a + PPT_x] - [RT_6 + (PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa}) \\
&\quad +(TD_d + PD)N_b] + T_{sw}, 0\}BW_{vci} \\
&= max\{[(TD_c + PD + PPT_x)(N_h - 2N_b) + (TD_d + PD)(N_a - N_b) + T_{sw} - (4PPT_{cwa} + 2TD_{wlc} \\
&\quad +2PD_{wl} + PPT_{cac})], 0\}BW_{vci}
\end{aligned}$$

Again, in most cases, since $RT_{BHO} < RT_{n1}$, $B_{CCN} = 0$.

B.3 Intra-Cluster Downlink With-Hint Analysis

$T_1 = PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	$RT_1 = T_1$
$T_2 = PPT_{cac} + (TD_c + PD + PPT_x)N_h + PPT_{cac}$	$RT_2 = RT_1 + T_2$
$T_3 = (TD_c + PD + PPT_x)N_h$	$RT_3 = RT_2 + T_3$
$T_4 = (TD_c + PD + PPT_x)N_a$	$RT_4 = RT_1 + T_4$
$T_5 = T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	$RT_5 = T_5$
$T_6 = (TD_c + PD + PPT_x)N_b + PPT_{cac}$	$RT_6 = min\{RT_2, RT_5\} + T_6$
$T_7 = (TD_c + PD + PPT_x)N_b$	$RT_7 = RT_6 + T_7$
$T_8 = PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	$RT_8 = min\{RT_2, RT_7\} + T_8$

Connection disruption time at MH:

$$\begin{aligned}
T_{disrupt} &= RT_8 \\
&= T_8 + min\{RT_2, RT_7\} \\
&= T_8 + min\{RT_2, (T_7 + min\{RT_2, RT_5\})\}
\end{aligned}$$

$$
\begin{aligned}
&= T_8 + min\{RT_2, (T_7 + RT_5)\} \\
&= T_8 + min\{(T_2 + T_1), (T_7 + T_5)\} \\
&= T_8 + min\{[2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 2PPT_{cac} + N_h(TD_c + PD + PPT_x)], [T_{acq} \\
&\quad +2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + PPT_{cac} + N_b(TD_c + PD + PPT_x)]\} \\
&= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + PPT_{cac} + PPT_x + min\{[PPT_{cac} + N_h(TD_c + PD \\
&\quad +PPT_x)], [T_{acq} + N_b(TD_c + PD + PPT_x)]\}
\end{aligned}
$$

Buffer requirement at CBS:

$$
\begin{aligned}
B_{CBS} &= RT_3 BW_{vci} \\
&= (T_3 + T_2 + T_1) BW_{vci} \\
&= [2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 2PPT_{cac} + 2N_h(TD_c + PD + PPT_x)] BW_{vci}
\end{aligned}
$$

Buffer requirement at NBS:

$$
\begin{aligned}
T_{BHO} &= (TD_d + PD)(N_a + N_h) + PPT_x \\
T_{n1} &= (TD_d + PD) N_b \\
B_{NBS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\} BW_{vci} \\
&= max\{(TD_d + PD)(N_a + N_h - N_b) + PPT_x + T_{sw}, 0\} BW_{vci}
\end{aligned}
$$

B.4 Intra-Cluster Downlink Without-Hint Analysis

T_1	=	$T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	RT_1	=	T_1
T_2	=	$PPT_{cac} + (TD_c + PD + PPT_x)N_h + PPT_{cac}$	RT_2	=	$RT_1 + T_2$
T_3	=	$(TD_c + PD + PPT_x)N_h$	RT_3	=	$RT_2 + T_3$
T_4	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_4	=	$RT_3 + T_4$
T_5	=	$(TD_c + PD + PPT_x)N_a$	RT_5	=	$RT_2 + T_5$
T_6	=	$(TD_c + PD + PPT_x)N_b + PPT_{cac}$	RT_6	=	$RT_1 + T_6$
T_7	=	$(TD_c + PD + PPT_x)N_b$	RT_7	=	$RT_6 + T_7$

Connection disruption time at MH:

$$\begin{aligned} T_{disrupt} &= RT_4 \\ &= (T_4 + T_3 + T_2 + T_1) \\ &= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 3PPT_{cac} + T_{acq} + 2N_h(TD_c + PD + PPT_x) \end{aligned}$$

Buffer requirement at CBS:

$$\begin{aligned} B_{CBS} &= [RT_2 + 2(TD_c + PD + PPT_x)]BW_{vci} \\ &= [T_2 + T_1 + 2(TD_c + PD + PPT_x)]BW_{vci} \\ &= [2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 3PPT_{cac} + T_{acq} + (N_h + 2)(TD_c + PD + PPT_x)]BW_{vci} \end{aligned}$$

Buffer requirement at NBS:

$$\begin{aligned} T_{BHO} &= (TD_d + PD)(N_a + N_h) + PPT_x \\ T_{n1} &= (TD_d + PD)N_b \\ B_{NBS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci} \\ &= max\{(TD_d + PD)(N_a + N_h - N_b) + PPT_x + T_{sw}, 0\}BW_{vci} \end{aligned}$$

B.5 Inter-Cluster Uplink With-Hint Analysis

T_1	$=$	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_1	$=$	T_1
T_2	$=$	$(TD_c + PD + PPT_x)N_h$	RT_2	$=$	$RT_1 + T_2$
T_3	$=$	$(TD_c + PD + PPT_x)N_a$	RT_3	$=$	$RT_1 + T_3$
T_4	$=$	$(TD_c + PD + PPT_x)N_{ax}$	RT_4	$=$	$RT_1 + T_4$
T_5	$=$	$T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	RT_5	$=$	T_5
T_6	$=$	$(TD_c + PD + PPT_x)N_b + PPT_{cac}$	RT_6	$=$	$min\{RT_2, RT_5\} + T_6$
T_7	$=$	$(TD_c + PD + PPT_x)N_{bx} + PPT_{cac}$	RT_7	$=$	$RT_6 + T_7$
T_8	$=$	$(TD_c + PD + PPT_x)N_{bx}$	RT_8	$=$	$RT_7 + T_8$
T_9	$=$	$(TD_c + PD + PPT_x)N_b$	RT_6	$=$	$RT_5 + T_9$
T_{10}	$=$	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_{10}	$=$	$RT_9 + T_{10}$

Connection disruption time at MH:

$$
\begin{aligned}
Tdisrupt &= RT_{10} \\
&= T_{10} + T_9 + T_8 + T_7 + T_6 + min\{(T_2 + T_1), T_5\} \\
&= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 2PPT_{cac} + 2(N_b + N_{bx})(TD_c + PD + PPT_x) \\
&\quad + min\{N_h(TD_c + PD + PPT_x), (T_{acq} + PPT_{cac})\}
\end{aligned}
$$

Buffer requirement at MGS:

$$
\begin{aligned}
RT_{BHO} &= RT_1 + (TD_d + PD)(N_a + N_{ax}) + PPT_x \\
RT_{n1} &= RT_{10} + (PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa}) + (TD_d + PD)(N_b + N_{bx}) \\
B_{CCN} &= max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\}BW_{vci} \\
&= max\{[(RT_1 - RT_{10}) + (TD_d + PD)(N_a + N_{ax} - N_b - N_{bx}) + PPT_x + T_{acq} - (2PPT_{cwa} + TD_{wlc} \\
&\quad + PD_{wl})], 0\}BW_{vci} \\
&= max\{[T_1 - (T_{10} + T_9 + T_8 + T_7 + T_6 + min\{RT_2, RT_5\}) + (TD_d + PD)(N_a + N_{ax} - N_b \\
&\quad - N_{bx}) + PPT_x + T_{sw} - (2PPT_{cwa} + TD_{wlc} + PD_{wl})], 0\}BW_{vci} \\
&= max\{[T_1 - (T_{10} + T_9 + T_8 + T_7 + T_6) + (TD_d + PD)(N_a + N_{ax} - N_b - N_{bx}) + PPT_x
\end{aligned}
$$

$$+T_{sw} - min\{(T_2 + T_1), T_5\} - (2PPT_{cwa} + TD_{wlc} + PD_{wl})], 0\}BW_{vci}$$

$$= max\{[T_{sw} + (TD_D + PD)(N_a + N_{ax} - N_b - N_{bx}) - (4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_{cac})$$

$$-2(N_b + N_{bx})(TD_c + PD + PPT_x) - min\{N_h(TD_c + PD + PPT_x), (T_{acq} + PPT_{cac})\}], 0\}BW_{vci}$$

Because $RT_{BHO} < RT_{n1}$ in general, $B_{CCN} = 0$.

B.6 Inter-Cluster Uplink Without-Hint Analysis

T_1	=	$T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	RT_1	=	T_1
T_2	=	$(TD_c + PD + PPT_x)N_h$	RT_2	=	$RT_1 + T_2$
T_3	=	$(TD_c + PD + PPT_x)N_a$	RT_3	=	$RT_2 + T_3$
T_4	=	$(TD_c + PD + PPT_x)N_{ax}$	RT_4	=	$RT_3 + T_4$
T_5	=	$(TD_c + PD + PPT_x)N_b + PPT_{cac}$	RT_5	=	$RT_1 + T_5$
T_6	=	$(TD_c + PD + PPT_x)N_{bx} + PPT_{cac}$	RT_6	=	$RT_5 + T_6$
T_7	=	$(TD_c + PD + PPT_x)N_{bx}$	RT_7	=	$RT_6 + T_7$
T_8	=	$(TD_c + PD + PPT_x)N_b$	RT_8	=	$RT_7 + T_8$
T_9	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_9	=	$RT_8 + T_9$

Connection disruption time at MH:

$$Tdisrupt = RT_9$$

$$= T_9 + T_8 + T_7 + T_6 + T_5 + T_1$$

$$= T_{acq} + 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 3PPT_{cac} + 2(N_b + N_{bx})(TD_c + PD + PPT_x)$$

Buffer requirement at MGS:

$$RT_{BHO} = RT_2 + (TD_d + PD)(N_a + N_{ax}) + PPT_x$$

$$RT_{n1} = RT_9 + (PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa}) + (TD_d + PD)(N_b + N_{bx})$$

$$B_{MGS} = max\{(RT_{BHO} - RT_{n1} + T_{sw}), 0\}BW_{vci}$$

$$= max\{[RT_2 - RT_9] + [(TD_d + PD)(N_a + N_{ax}) + PPT_x - (TD_d + PD)(N_b + N_{bx}) - (2PPT_{cwa}$$

$$
\begin{aligned}
& +TD_{wlc} + PD_{wl})] + T_{sw}, 0\}BW_{vci} \\
= \;& max\{[T_2 + T_1 - (T_9 + T_8 + T_7 + T_6 + T_5 + T_1)] + [(TD_d + PD)(N_a + N_{ax} - N_b - N_{bx}) \\
& -(2PPT_{cwa} + TD_{wlc} + PD_{wl})] + T_{sw}, 0\}BW_{vci} \\
= \;& max\{[T_{sw} + (TD_d + PD)(N_a + N_{ax} - N_b - N_{bx}) - (TD_c + PD + PPT_x)(N_h - 2N_b - 2N_{bx}) \\
& -(2PPT_{cwa} + TD_{wlc} + PD_{wl})], 0\}BW_{vci}
\end{aligned}
$$

Usually $RT_{BHO} < RT_{n1}$, so $B_{CCN} = 0$.

B.7 Inter-Cluster Downlink With-Hint Analysis

T_1	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_1	=	T_1
T_2	=	$PPT_{cac} + (TD_c + PD + PPT_x)N_h + PPT_{cac}$	RT_2	=	$RT_1 + T_2$
T_3	=	$(TD_c + PD + PPT_x)N_h$	RT_3	=	$RT_2 + T_3$
T_4	=	$(TD_c + PD + PPT_x)N_a$	RT_4	=	$RT_1 + T_4$
T_5	=	$(TD_c + PD + PPT_x)N_{ax}$	RT_5	=	$RT_4 + T_5$
T_6	=	$T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	RT_6	=	T_6
T_7	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_7	=	$min\{RT_2, RT_{11}\} + T_7$
T_8	=	$(TD_c + PD + PPT_x)N_b + PPT_{cac}$	RT_8	=	$min\{RT_2, RT_6\} + T_8$
T_9	=	$(TD_c + PD + PPT_x)N_{bx} + PPT_{cac}$	RT_9	=	$RT_8 + T_9$
T_{10}	=	$(TD_c + PD + PPT_x)N_{bx}$	RT_{10}	=	$RT_9 + T_{10}$
T_{11}	=	$(TD_c + PD + PPT_x)N_b$	RT_{11}	=	$RT_{10} + T_{11}$

Connection disruption time at MH:

$$
\begin{aligned}
T_{disrupt} &= RT_7 \\
&= T_7 + min\{RT_2, RT_{11}\} \\
&= T_7 + min\{RT_2, (T_{11} + T_{10} + T_9 + T_8 + min\{RT_2, RT_6\})\}
\end{aligned}
$$

$$
\begin{aligned}
&= T_7 + min\{T_2 + T_1, (T_{11} + T_{10} + T_9 + T_8 + T_6)\} \\
&= T_7 + (2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x) + min\{N_h(TD_c + PD + PPT_x), T_{acq} + PPT_{cac} \\
&\quad +2(N_b + N_{bx})(TD_c + PD + PPT_x)\} \\
&= 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + min\{N_h(TD_c + PD + PPT_x), T_{acq} + PPT_{cac} \\
&\quad +2(N_b + N_{bx})(TD_c + PD + PPT_x)\}
\end{aligned}
$$

Buffer requirement at CBS:

$$
\begin{aligned}
B_{CBS} &= RT_3 BW_{vci} \\
&= (T_3 + T_2 + T_1)BW_{vci} \\
&= [(2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 2PPT_{cac}) + 2N_h(TD_c + PD + PPT_x)]BW_{vci}
\end{aligned}
$$

Buffer requirement at NBS:

$$
\begin{aligned}
T_{BHO} &= (TD_d + PD)(N_a + N_{ax} + N_h) + PPT_x \\
T_{n1} &= (TD_d + PD)(N_b + N_{bx}) \\
B_{NBS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci} \\
&= max\{(TD_d + PD)(N_a + N_{ax} + N_h - N_b - N_{bx}) + PPT_x + T_{sw}, 0\}BW_{vci}
\end{aligned}
$$

B.8 Inter-Cluster Downlink Without-Hint Analysis

T_1	=	$T_{acq} + PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x + PPT_{cac}$	RT_1	=	T_1
T_2	=	$PPT_{cac} + (TD_c + PD + PPT_x)N_h + PPT_{cac}$	RT_2	=	$RT_1 + T_2$
T_3	=	$(TD_c + PD + PPT_x)N_h$	RT_3	=	$RT_2 + T_3$
T_4	=	$PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_{cwa} + PPT_x$	RT_4	=	$RT_3 + T_4$
T_5	=	$(TD_c + PD + PPT_x)N_a$	RT_5	=	$RT_2 + T_5$
T_6	=	$(TD_c + PD + PPT_x)N_{ax}$	RT_6	=	$RT_5 + T_6$
T_7	=	$(TD_c + PD + PPT_x)N_b + PPT_{cac}$	RT_7	=	$RT_1 + T_7$
T_8	=	$(TD_c + PD + PPT_x)N_{bx} + PPT_{cac}$	RT_8	=	$RT_7 + T_8$
T_9	=	$(TD_c + PD + PPT_x)N_{bx}$	RT_9	=	$RT_8 + T_9$
T_{10}	=	$(TD_c + PD + PPT_x)N_b$	RT_{10}	=	$RT_9 + T_{10}$

Connection disruption time at MH:

$$\begin{aligned} T_{disrupt} &= RT_4 \\ &= T_4 + T_3 + T_2 + T_1 \\ &= T_{acq} + 4PPT_{cwa} + 2TD_{wlc} + 2PD_{wl} + 2PPT_x + 3PPT_{cac} + 2N_h(TD_c + PD + PPT_x) \end{aligned}$$

Buffer requirement at CBS:

$$\begin{aligned} B_{CBS} &= [RT_2 + 2(TD_c + PD + PPT_x)]BW_{vci} \\ &= [(T_2 + T_1) + 2(TD_c + PD + PPT_x)]BW_{vci} \\ &= [2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 3PPT_{cac} + (TD_c + PD + PPT_x)N_h + T_{acq} + 2(TD_c \\ &\quad + PD + PPT_x)]BW_{vci} \\ &= [2PPT_{cwa} + TD_{wlc} + PD_{wl} + PPT_x + 3PPT_{cac} + (N_h + 2)(TD_c + PD + PPT_x)]BW_{vci} \end{aligned}$$

Buffer requirement at NBS:

$$T_{BHO} = (TD_d + PD)(N_{ax} + N_a + N_h) + PPT_x$$
$$T_{n1} = (TD_d + PD)(N_{bx} + N_b)$$
$$B_{NBS} = max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci}$$
$$= max\{(TD_d + PD)(N_{ax} + N_a + N_h - N_{bx} - N_b) + PPT_x + T_{sw}, 0\}BW_{vci}$$

B.9 Inter-Domain Phase 2

$T_1 = (TD_c + PD + PPT_x)N_{ax}$	$RT_1 = T_1$
$T_2 = (TD_c + PD + PPT_x)N_{cx}$	$RT_2 = RT_1 + T_1$
$T_3 = PPT_{cac} + (TD_c + PD + PPT_x)N_{bx} + PPT_{cac}$	$RT_3 = T_3$
$T_4 = (TD_c + PD + PPT_x + PPT_{cac})N_{dx}$	$RT_4 = RT_3 + T_4$
$T_5 = (TD_c + PD + PPT_x)N_{dx}$	$RT_5 = RT_4 + T_5$
$T_6 = (TD_c + PD + PPT_x)N_{bx}$	$RT_6 = RT_5 + T_6$

B.9.1 Uplink Case

Buffer requirement at CCN:

$$T_{BHO} = (TD_d + PD)(N_{ax} + N_{cx}) + PPT_x$$
$$T_{n1} = (TD_d + PD)(N_{bx} + N_{dx})$$
$$B_{CCN} = max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci}$$
$$= max\{(TD_d + PD)(N_{ax} + N_{cx} - N_{bx} - N_{dx}) + PPT_x + T_{sw}, 0\}BW_{vci}$$

B.9.2 Downlink Case

Buffer requirement at MCS:

$$T_{BHO} = (TD_d + PD)(N_{cx} + N_{ax}) + PPT_x$$

$$
\begin{aligned}
T_{n1} &= (TD_d + PD)(N_{dx} + N_{bx}) \\
B_{MCS} &= max\{(T_{BHO} - T_{n1} + T_{sw}), 0\}BW_{vci} \\
&= max\{(TD_d + PD)(N_{ax} + N_{cx} - N_{bx} - N_{dx}) + PPT_x + T_{sw}, 0\}BW_{vci}
\end{aligned}
$$

Bibliography

Bibliography

[1] T. S. Rappaport, *Wireless Communications: Principles & Practice.* Prentice-Hall, Inc., 1996.

[2] N. J. Muller, *Wireless Data Networking.* Artech House Publishers, 1995.

[3] D. C. Cox, "Personal communication—a viewpoint," *IEEE Communications Magazine*, vol. 28, November 1994.

[4] D. C. Cox, "Wireless network access for personal communications," *IEEE Communications Magazine*, vol. 31, Dec. 1992.

[5] D. C. Cox, "A radio system proposal for widespread low-power tetherless communication," *IEEE Communications Magazine*, vol. 29, Feb. 1991.

[6] D. Goodman, "Cellular packet communications," *IEEE Transaction on Communications*, vol. 28, Aug. 1990.

[7] D. Goodman, "Trends in cellular and cordless communications," *IEEE Communications Magazine*, Feb. 1991.

[8] I. Ross, "Wireless network directions," *IEEE Communications Magazine*, vol. 29, June 1991.

[9] R. Steele, "The evolution of personal communications," *IEEE Personal Communications Magazine*, pp. 6–11, 1994.

[10] W. Stallings, *ISDN and Broadband ISDN.* MacMillan Publishing Company, 2 ed., 1992.

[11] S. E. Minzer, "Broadband ISDN and asynchronous transfer mode," *IEEE Communications Magazine*, pp. 17–24, September 1989.

[12] A. Hac and H. B. Mutlu, "Synchronous optical network and Broadband ISDN protocols," *IEEE Computer Magazine*, vol. 11, pp. 26–34, November 1989.

[13] The ATM Forum, *ATM USER-NETWORK INTERFACE (UNI) SPECIFICATION*, 3.1 ed., 1995.

[14] D. Raychaudhuri, "ATM-based transport architecture for multiservices wireless personal communication networks," in *Proceedings of IEEE International Conference on Communications*, pp. 559–565, 1994.

[15] A. Acampora, "Wireless ATM: A perspective on issues and prospects," *IEEE Personal Communication Magazine*, pp. 8–17, August 1996.

[16] 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, January 1997.

[17] S. Aikawa, Y. Motoyama, and M. Umehira, "Forward error correction schemes for wireless ATM systems," in *Proceedings of IEEE International Conference on Communications*, pp. 454–458, 1996.

[18] L. L. Peterson and B. S. Davie, *Computer Networks.* Morgan Kaufmann Publishers, Inc., 1996.

[19] A. Tanenbaum, *Computer Networks.* Addison Wisley Publishers, Inc., 1996.

[20] E. Ayanoglu, K. Y. Eng, and M. J. Karol, "Wireless ATM: Limits, challenges, and proposals," *IEEE Personal Communications Magazine*, pp. 18–34, August 1996.

[21] E. I. A. I. Association, "IS-41 cellular radio-telecommunications intersystems operations," January 1993.

[22] I. Katzela and M. Veeraraghavan, "Virtual trees routing protocol for a wireless ATM LAN," in *Proceedings of IEEE International Conference on Universal Personal Communications*, pp. 369–373, September 1996.

[23] C.-K. Toh, "The design and implementation of a hybrid handover protocol for multi-media wireless LANs," in *Proceedings of ACM First International Conference on Mobile Computing and Networking (MOBICOM' 95)*, November 1995.

[24] A. S. Acampora and M. Naghshineh, "An architecture and methodology for mobile-executed handoff in cellular ATM networks," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 1365–1375, October 1994.

[25] B. A. Akyol and D. C. Cox, "Rerouting for handoff in a wireless ATM network," in *Proceedings of IEEE International Conference on Universal Personal Communications*, pp. 374–379, 1996.

[26] R. Yuan, S. K. Biswas, and D. Raychaudhuri, "A signaling and control architecture for mobility support in wireless ATM networks," in *Proceedings of IEEE International Conference on Communications*, pp. 478–483, 1996.

[27] R. Yuan, S. K. Biswas, and D. Raychaudhuri, "B-ISDN architectures and protocols to support wireless personal communications internetworking," in *Proceedings of IEEE Conference on Personal Indoor Mobile Radio Communications*, (Toronto, Canada), pp. 27–29, September 1995.

[28] J. Gardener and B. West, *Personal Communication Systems and Technologies.* Artech House, 1994.

[29] D. Raychaudhuri and N. D. Wilson, "ATM-based transport architecture for multiservices wireless personal communication networks," *IEEE Journal on Selected Areas in Communications*, pp. 1401–1414, 1992.

[30] I. Brodsky, *Wireless: The Revolution in Personal Telecommunications.* Artech House Publishers, 1995.

[31] C. G. Omidyar and A. Aldridge, "Introduction to SDH/SONET," *IEEE Communications Magazine*, pp. 30–33, Sept. 1993.

[32] Y. Yeh, M. G. Hluchj, and A. S. Acampora, "The Knockout switch: A simple, modular architecture for high-performance packet switching," *IEEE Journal on Selected Areas of Communications*, vol. SAC-5, pp. 1274–1282, October 1987.

[33] N. Giacopelli, "Sunshine: A high-performance self-routing broadband packet switch architecture," *IEEE Journal on Selected Areas of Communications*, vol. 9, pp. 1289–1298, October 1991.

[34] H.-T. Lin and H. D. Hughes, "Performance evaluation of priority output buffers in ATM switches," in *Proceedings of IEEE International Conference on Computer, Communication, and Networks*, pp. 83–87, September 1994.

[35] H.-T. Lin and H. D. Hughes, "Performance evaluation of priority buffer management in ATM networks," *Journal of Computer System and System Engineering*, 1996. in press.

[36] Q. Ta and J. S. Meditch, "A high speed integrated services switch based on 4 x 4 switch element," in *IEEE INFOCOM*, pp. 1164–1171, June 1990.

[37] H. Ahmadi, "A high-performance switch fabric for integrated circuit and packet switching," in *IEEE INFOCOM*, pp. 9–18, March 1988.

[38] K. E. Batcher, "Sorting networks and their applications," in *Sprtng. Joint Computer Conference*, pp. 307–314, 1968.

[39] C. Clos, "A study of non-blocking switching networks," *Bell System Technical Journal*, vol. 32, pp. 406–424, 1953.

[40] "B-ISDN and how it works," *IEEE Spectrum Magazine*, pp. 39–44, August 1991.

[41] J. Kurose, "Open issues and challenges in providing quality of service guarantees in high-speed networks," in *Proceedings of ACM SIGCOMM*, pp. 6–15, 1991.

[42] G. Ramamurthy and R. S. Dighe, "A multidimensional framework for congestion control in B-ISDN," *IEEE Journal on Selected Areas in Communications*, December 1991.

[43] P. W. Tse and M. Zukerman, "Connection admission control in ATM networks," in *Proceedings of IEEE Globe Telecommunications Conference*, pp. 1790–1794, November 1994.

[44] K. Patel, B. C. Smith, and L. A. Rowe, "Performance of a software MPEG video decorder," in *Proceedings of ACM Multimedia Conference*, 1993.

[45] D. L. Gall, "MPEG: A video compression standard for multimedia applications," *Communications of ACM*, vol. 34, pp. 47–58, April 1991.

[46] P. Pancha and M. E. Zarki, "MPEG coding for variable bit rate video transmission," *IEEE Communications Magazine*, vol. 32, pp. 54–66, May 1994.

[47] D. Reininger and et al., "Statistical multiplexing of VBR MPEG compressed video on ATM networks," in *Proceedings of IEEE INFOCOM*, vol. 3, pp. 919–926, 1993.

[48] P. S. The ATM Forum, "ATM forum 94-0471r13 PNNI draft specification," 1994.

[49] M. Mouly and M. B. Pautet, "The GSM system for mobile communications." 49 rue Louise Bruneau, Palaiseau, France, 1992.

[50] Y. B. Lin and S. K. DeVries, "PCS network signaling using SS7," *IEEE Communications Magazine*, pp. 44–55, June 1995.

[51] A. R. Modarressi and R. A. Skoog, "Signaling system 7: A tutorial," *IEEE Communications Magazine*, vol. 28, pp. 19–35, July 1990.

[52] D. R. Wilson, "Signaling system no. 7, IS-41 and cellular telephony networking," in *Proceedings of IEEE*, pp. 664–652, April 1992.

[53] S. Mohan and R. Jain, "Two user location strategies for personal communication services," *IEEE Personal Communications Magazine*, pp. 42–50, first quarter, 1994.

[54] I. F. Akyildiz and J. S. Ho, "On location management for Personal Communications Networks," *IEEE Communications Magazine*, pp. 138–145, September 1996.

[55] C. N. Lo, S. Mohan, and R. S. Wolff, "Performance modeling and simulation of data management for personal communication applications," in *Proceedings of IEEE Conference on Personal Indoor Mobile Radio Communications*, November 1992.

[56] C. N. Lo, R. S. Wolff, and R. C. Bernhardt, "An estimate of network database transaction volume to support personal communications services," in *Proceedings of IEEE International Conference on Universal Personal Communications*, November 1992.

[57] R. Jain, Y. B. Lin, and S. Mohan, "A caching strategy to reduce network impacts of PCS," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 1434–1444, October 1994.

[58] N. Shivakumar and J. Widom, "User profile replication for faster location lookup in mobile environments," in *Proceedings of ACM First International Conference on Mobile Computing & Networking (MOBICOM' 95)*, pp. 161–169, November 1995.

[59] R. Jain and Y. B. Lin, "An auxiliary user location strategy employing forwarding pointers to educe network impact of PCS," *ACM-Baltzer Journal of Wireless Networks*, vol. 1, pp. 197–210, July 1995.

[60] J. S. M. Ho and I. F. Akyildiz, "Local anchor scheme for reducing location tracking costs in PCN," *IEEE/ACM Transaction on Networking*, 1996.

[61] J. Z. Wang, "A fully distributed location registration strategy for universal personal communication systems," *IEEE Journal of Selected Areas on Communications*, vol. 11, pp. 850–860, August 1993.

[62] V. A. et. al., "Optimization of a database hierarchy for mobility tracking in a personal communications network," *Performance Evaluation*, vol. 20, pp. 287–300, May 1994.

[63] B. R. Badrinath, T. Imielinski, and A. Virmani, "Locating strategies for personal communication network," in *Proceedings of Workshop on Networking of Personal Communication Applications*, December 1992.

[64] K. Y. Eng, M. J. Karol, M. Veeraraghavan, E. Ayanoglu, C. B. Woodworth, P. Pancha, and R. A. Valenzuela, "BAHAMA: A broadband ad-hoc wireless ATM local-area network," in *Proceedings of IEEE International Conference on Communications*, pp. 1216–1223, June 1995.

[65] K. Y. Eng, M. J. Karol, M. Veeraraghavan, E. Ayanoglu, C. B. Woodworth, P. Pancha, and R. A. Valenzuela, "A wireless broadband ad-hoc ATM local-area network," in *ACM Journal of Wireless Networks*, pp. 161–174, 1995.

[66] A. S. Acampora and M. Naghshineh, "Control and quality-of-service provisioning in high-speed microcellular networks," *IEEE Personal Communications Magazine*, pp. 36–43, 1994.

[67] B. A. Akyol and D. C. Cox, "Handling mobility in a wireless ATM network," in *Proceedings of IEEE INFOCOM*, pp. 1405–1413, May 1996.

[68] J. Porter and A. Hopper, "An ATM-based protocol for wireless LANs," Tech. Rep. 94.2, Olivetti Research Ltd., April 1994. Available at ftp://ftp.cam-ori.co.uk/pub/docs/ORL/tr.94.2.ps.Z.

[69] P. Agrawal, E. Hyden, P. Krzyzanowski, P. Mishra, M. B. Srivastava, and J. A. Trotter, "SWAN: A mobile multimedia wireless network," *IEEE Personal Communications Magazine*, 1996.

[70] D. Raychaudhuri, "Architecture, system design and prototyping of a wireless ATM network," in *Proceedings of IEEE ATM workshop*, November 1995.

[71] D. Raychaudhuri, "Wireless ATM network: Architecture, system design and prototyping," *IEEE Personal Communications Magazine*, pp. 42–49, August 1996.

[72] K. Keeton, B. A. Mah, S. Seshan, R. H. Katz, and D. Ferrari, "Providing connection-oriented network services to mobile hosts," in *Proceedings of the USENIX Symposium on Mobile and Location-Independent Computing*, August 1993.

[73] H.-T. Lin and H. D. Hughes, "A handoff scheme for wireless ATM networks," in *Proceedings of International Conference on System Engineering*, pp. 509–514, July 1996.

[74] H.-T. Lin and H. D. Hughes, "An experimental handoff scheme for wireless ATM networks," in *Proceedings of IEEE International Conference on Universal Personal Communication*, pp. 397–401, September 1996.

[75] W. C. Lee, "Topology aggregation for hierarchical routing in ATM networks," in *Proceedings of ACM SIGCOMM*, pp. 82–92, 1995.

[76] J. S. M. Ho and I. F. Akyildiz, "Dynamic hierarchical database architecture for location management in PCS networks," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 646–660, October 1997.

[77] M. Veeraraghavan, P. Pancha, and G. Dommety, "Connectionless ATM using an ATM switch router," in *Proceedings of ECMAST*, May 1997.

[78] R. Want, A. Hopper, and et. al., "The active badge location system," *ACM Transactions on Information Systems*, vol. 10, pp. 91–102, January 1992.

[79] C.-K. Toh, "Performance evaluation of crossover switch discovery algorithms for wireless ATM LANs," in *Proceedings of IEEE INFOCOM*, pp. 1380–1387, 1996.

[80] G. P. Polline and K. S. Meier-Hellstern, "Efficient routing of information between interconnected cellular mobile switching centers," *IEEE/ACM Transactions of Networking*, vol. 3, pp. 765–774, December 1995.

[81] D. Raychaudhuri, L. J. French, R. J. Siracusa, S. K. Biswas, R. Yuan, P. Narasimhasn, and C. A. Johnston, "WATMnet: A prototype wireless ATM system for multimedia personal communication," *IEEE Journal on Selected Areas in Communications*, pp. 83–94, January 1997.

[82] M. Selsh and L. Kaufman, *Running Linux*. O'Reilly Publishers, 2 ed., 1996.

[83] M. Laubach, "Classical IP and ARP over ATM." http://ds.internic.net/rfc/rfc1577.txt, January 1994.

[84] W. Almesberger, "ATM on Linux user's guide, release 0.31." http://lrcwww.epfl.ch/linux-atm, 1997.

[85] W. Almesberger, "Linux ATM API draft, version 0.4." http://lrcwww.epfl.ch/linux-atm, July 1996.

[86] W. Almesberger, "ATM on Linux." http://lrcwww.epfl.ch/linux-atm, March 1996.

[87] W. R. Stevens, *TCP/IP Illustrated*, vol. 1. Addison Wisley Publishers, Inc., 1994.

[88] V. Jacobson, "Congestion avoidance and control," in *SIGCOMM*, August 1988.

[89] E. Amir, H. Balakrishnan, S. Seshan, and R. Katz, "Efficient TCP over networks with wireless links," in *Proceedings of Workshop on Hot Topics in Operating Systems*, May 1995.