On the Merits of Deploying TDM-based Next-Generation PON Solutions in the Access Arena As Multiservice, All Packet-Based 4G Mobile Backhaul RAN Architecture

Syed Rashid N. Zaidi
Graduate Center, City University of New York
On the Merits of Deploying TDM-based Next-Generation PON Solutions in
the Access Arena As Multiservice, All Packet-Based 4G Mobile Backhaul
RAN Architecture

By

Syed Rashid Nafees Zaidi

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2014
© 2014

Syed Rashid Nafees Zaidi

All Rights Reserved
This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Prof. Mohamed A. Ali

Date Chair of Examining Committee

Dean Ardie D. Walser

Date Executive Officer

Prof. Roger Dorsinville, EE Dept, CCNY, CUNY

Prof. Barry Gross, EE Dept, CCNY, CUNY

Prof. Neophytos Antoniades, Engr. Dept., College of Staten Island, CUNY

Dr. Andrew Wallace, AT&T

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

On the Merits of Deploying TDM-based Next-Generation PON Solutions in the Access Arena As Multiservice, All Packet-Based 4G Mobile Backhaul RAN Architecture

By

Syed Rashid Nafees Zaidi

Adviser: Professor Mohamed Ali

The phenomenal growth of mobile backhaul capacity required to support the emerging fourth-generation (4G) traffic including mobile WiMAX, cellular Long-Term Evolution (LTE), and LTE-Advanced (LTE-A) requires rapid migration from today’s legacy circuit switched T1/E1 wireline and microwave backhaul technologies to a new fiber-supported, all-packet-based mobile backhaul infrastructure. Clearly, a cost effective fiber supported all-packet-based mobile backhaul radio access network (RAN) architecture that is compatible with these inherently distributed 4G RAN architectures is needed to efficiently scale current mobile backhaul networks. However, deploying a green fiber-based mobile backhaul infrastructure is a costly proposition mainly due to the significant cost associated with digging the trenches in which the fiber is to be laid.

These, along with the inevitable trend towards all-IP/Ethernet transport protocols and packet switched networks, have prompted many carriers around the world to consider the potential of
utilizing the existing fiber-based Passive Optical Network (PON) access infrastructure as an all-packet-based converged fixed-mobile optical access networking transport architecture to backhaul both mobile and typical wireline traffic. Passive Optical Network (PON)-based fiber-to-the-curb/home (FTTC/FTTH) access networks are being deployed around the globe based on two Time-Division Multiplexed (TDM) standards: ITU G.984 Gigabit PON (GPON) and IEEE 802.ah Ethernet PON (EPON). A PON connects a group of Optical Network Units (ONUs) located at the subscriber premises to an Optical Line Terminal (OLT) located at the service provider’s facility.

It is the purpose of this thesis to examine the technological requirements and assess the performance analysis and feasibility for deploying TDM-based next-generation (NG) PON solutions in the access arena as multiservice, all packet-based 4G mobile backhaul RAN and/or converged fixed-mobile optical networking architecture. Specifically, this work proposes and devises a simple and cost-effective 10G-EPON-based 4G mobile backhaul RAN architecture that efficiently transports and supports a wide range of existing and emerging fixed-mobile advanced multimedia applications and services along with the diverse quality of service (QoS), rate, and reliability requirements set by these services. The techno-economics merits of utilizing PON-based 4G RAN architecture versus that of traditional 4G (mobile WiMAX and LTE) RAN will be thoroughly examine and quantified.

To achieve our objective, we utilize the existing fiber-based PON access infrastructure with novel ring-based distribution access network and wireless-enabled OLT and ONUs as the multiservice packet-based 4G mobile backhaul RAN infrastructure. Specifically, to simplify the implementation of such a complex undertaking, this work is divided into two sequential phases.
In the first phase, we examine and quantify the overall performance of the standalone ring-based 10G-EPON architecture (just the wireline part without overlaying/incorporating the wireless part (4G RAN)) via modeling and simulations. We then assemble the basic building blocks, components, and sub-systems required to build up a proof-of-concept prototype testbed for the standalone ring-based EPON architecture. The testbed will be used to verify and demonstrate the performance of the standalone architecture, specifically, in terms of power budget, scalability, and reach.

In the second phase, we develop an integrated framework for the efficient interworking between the two wireline PON and 4G mobile access technologies, particularly, in terms of unified network control and management (NCM) operations. Specifically, we address the key technical challenges associated with tailoring a typically centralized PON-based access architecture to interwork with and support a distributed 4G RAN architecture and associated radio NCM operations. This is achieved via introducing and developing several salient-networking innovations that collectively enable the standalone EPON architecture to support a fully distributed 4G mobile backhaul RAN and/or a truly unified NG-PON-4G access networking architecture. These include a fully distributed control plane that enables intercommunication among the access nodes (ONUs/BSs) as well as signaling, scheduling algorithms, and handoff procedures that operate in a distributed manner.

Overall, the proposed NG-PON architecture constitutes a complete networking paradigm shift from the typically centralized PON’s architecture and OLT-based NCM operations to a new disruptive fully distributed PON’s architecture and NCM operations in which all the typically centralized OLT-based PON’s NCM operations are migrated to and independently implemented
by the access nodes (ONUs) in a distributed manner. This requires migrating most of the
typically centralized wireline and radio control and user-plane functionalities such as dynamic
bandwidth allocation (DBA), queue management and packet scheduling, handover control, radio
resource management, admission control, etc., typically implemented in today’s OLT/RNC, to
the access nodes (ONUs/4G BSs).

It is shown that the overall performance of the proposed EPON-based 4G backhaul including
both the RAN and Mobile Packet Core (MPC) (Evolved Packet Core (EPC) per 3GPP LTE’s
standard) is significantly augmented compared to that of the typical 4G RAN, specifically, in
terms of handoff capability, signaling overhead, overall network throughput and latency, and
QoS support. Furthermore, the proposed architecture enables redistributing some of the
intelligence and NCM operations currently centralized in the MPC platform out into the access
nodes of the mobile RAN. Specifically, as this work will show, it enables offloading sizable
fraction of the mobile signaling as well as actual local upstream traffic transport and processing
(LTE bearers switch/set-up, retain, and tear-down and associated signaling commands from the
BSs to the EPC and vice-versa) from the EPC to the access nodes (ONUs/BSs). This has a
significant impact on the performance of the EPC. First, it frees up a sizable fraction of the badly
needed network resources as well as processing on the overloaded centralized serving nodes
(AGW) in the MPC. Second, it frees up capacity and sessions on the typically congested mobile
backhaul from the BSs to the EPC and vice-versa.
Acknowledgements

My thesis could not have been finished without the support and cooperation of many people.

I would first like to express my sincere appreciation and gratitude to Professor Mohammad Ali for his support, guidance, discussions and exhaustive editing of my publications and thesis. Professor Mohamed A. Ali has been a source of enthusiasm and encouragement over many years. His guidance, discussions and exhaustive editing of my publications and thesis make this work possible. Without his support, I could not imagine the completion of my PhD at Graduate Center, CUNY.

I also like to extend my gratitude to Prof. Roger Dorsinville who is inspiration for me for various research publications. I am also grateful to Prof. Neophytos Antoniades for his valuable time and guidance during the course of research work. Moreover I am also thankful to Prof. Barry Gross and Dr. Andrew Wallace for serving on my supervisory committee. Their insightful recommendations have improved the quality of this work.

Much appreciation is given to my colleague and friend Shahab Hussain for all the technical discussions and cooperation and for all the valuable and productive time we shared.
My dissertation is dedicated to my parents, siblings, my wife and my children who have been constantly supportive of me during my years of life, growth and education.
# TABLE OF CONTENTS

1 INTRODUCTION .............................................................................................................. 1

1.1 Overview ...................................................................................................................... 2

1.2 Thesis Motivation.......................................................................................................... 5

1.3 Thesis Statement .......................................................................................................... 9

1.4 Thesis Organization ..................................................................................................... 13

2 OVERVIEW OF 4G WIRELESS TECHNOLOGIES-WIMAX & LTE ............. 15

2.1 Introduction .................................................................................................................. 16

2.2 LONG TERM EVOLUTION (LTE) .................................................................................. 17
  2.2.1 LTE Basics .............................................................................................................. 17
  2.2.2 Physical Parameters and Frame Structure ................................................................. 21
  2.2.3 MIMO Transmissions .......................................................................................... 24

2.3 WORLDWIDE INTEROPERABILITY FOR MICROWAVE ACCESS (WIMAX) ........... 24
  2.3.1 WiMAX Network Architecture ............................................................................ 24
  2.3.2 IEEE 802.16-2005 Air Interface ........................................................................... 26
  2.3.3 MIMO Transmissions .......................................................................................... 27

2.4 COMPARISON OF NEXT GENERATION MOBILE WIRELESS TECHNOLOGIES, LTE AND WIMAX ................................................................................................................. 28

3 FULLY DISTRIBUTED RING-BASED 10G-EPON ARCHITECTURE ....... 31

3.1 Introduction .................................................................................................................. 32
3.2 Standalone Ring-Based 10-G-EPON Architecture

3.2.1 Normal State Operation

3.2.2 Protected State Architecture

3.3 Recovery Time

3.4 Power Budget and scalability Analysis

3.5 Experiment

4 EPON-BASED 4G MOBILE BACKHAUL RAN ARCHITECTURE

4.1 Introduction

4.2 EPON-Based 4G Mobile Backhaul RAN Architecture and/or Converged Fixed Mobile Optical Access Network Architecture

5 KEY ENABLING NETWORKING INNOVATIONS

5.1 Fully Distributed Control Plane

5.2 Fully Distributed Dynamic Bandwidth Allocation at the ONUs/BSs

5.2.1 Overview of Typical PON Scheduling Schemes

5.2.2 Decentralized Dynamic Bandwidth Allocation Scheme

5.3 QoS Mapping

5.4 Layer-2 versus Layer-3 Connectivity at the Transport Layer

6 KEY TECHNOLOGICAL AND ECONOMIC ADVANTAGES ENABLED BY THE DISTRIBUTED EPON-BASED RAN ARCHITECTURE
6.1 Enhanced Network Throughput and Stability .................................................. 72

6.2 Significance of Local Mobile LAN Traffic: ..................................................... 78

6.3 Mobility Management - Enhanced Handoff Capabilities.................................. 83
  6.3.1 IEEE 802.16e Handover Process Overview............................................. 85
  6.3.2 Proposed Handover Scheme ................................................................. 93

6.4 Illustrative Examples ....................................................................................... 97
  6.4.1 Handover Scenario 1 .............................................................................. 98
  6.4.2 More Handoff Scenarios......................................................................... 107
  6.4.3 Handover Scenario 2 ............................................................................ 108
  6.4.4 Handover Scenario 3 ............................................................................ 111
  6.4.5 Handover Scenario 4 ............................................................................ 115

6.5 Conclusion ...................................................................................................... 118

7 REFERENCES ................................................................................................... 120

7.1 Chapter 1 .................................................................................................... 120

7.2 Chapter 2 .................................................................................................... 122

7.3 Chapter 3 .................................................................................................... 123

7.4 Chapter 4 .................................................................................................... 123

7.5 Chapter 5 .................................................................................................... 124

7.6 Chapter 6 .................................................................................................... 126

8 PUBLICATIONS ............................................................................................... 128
**LIST OF FIGURES**

| Figure 2-1: LTE Network Architecture | ......................................................... 18 |
| Figure 2-2: Orthogonal sub-channels of OFDM system over bandwidth W | ................................. 20 |
| Figure 2-3: LTE Frame | ................................................................. 23 |
| Figure 2-4: LTE Downlink Physical Resource Block | ......................................................... 23 |
| Figure 2-5: WiMAX Network Architecture | ......................................................... 25 |
| Figure 2-6: OFDMA TDD Frame Structure | ......................................................... 27 |
| Figure 3-1: Standalone 10G-EPON Architecture | ......................................................... 35 |
| Figure 3-2: ONU Architecture | ................................................................. 36 |
| Figure 3-3: 10G EPON Protection State Architecture | ......................................................... 38 |
| Figure 3-4: 10 Nodes Network Architecture | ......................................................... 41 |
| Figure 3-5: 7 Nodes Protected State Network Architecture | ......................................................... 42 |
| Figure 3-6: Inside structures of (a) Distribution node (b) ONUs excluding the last one and (c) last ONU | ................................................................. 43 |
| Figure 3-7: Experimental setup for the measurement of BER of the ONU before last under 10Gbps downstream communication | ......................................................... 43 |
| Figure 3-8: Measured (a) bit error rate under different received power (b) eye diagram at -17.5dBm received power of the ONU before last | ......................................................... 44 |
| Figure 3-9: Measured bit error rate of same structure at 2.5Gbps | ......................................................... 45 |
| Figure 4-1: The proposed converged NG-PON1-LTE RAN access architecture | ......................................................... 48 |
| Figure 4-2: (a) Architecture of the ONU-eNB, (b) Functional Modules hardware layout | ................................. 51 |
Figure 4-3: (a) Architecture of the ONU-BS, (b) Functional Modules hardware layout …….. 52

Figure 5-1: (a) Architecture of the ONU-eNB, (b) Functional Modules hardware layout …….. 65

Figure 6-1: Uplink Utilization Time Series for Evenly loaded BSs ............................... 75

Figure 6-2: Uplink Utilization Time Series for Unevenly loaded BSs ............................. 76

Figure 6-3: Uplink Ave Throughput ........................................................................ 76

Figure 6-4: Average network ETE delay for rtPs and BE traffic .................................... 77

Figure 6-5: WiMAX Backhaul Uplink Utilization with evenly loaded BSs ...................... 77

Figure 6-6: WiMAX Backhaul Uplink Utilization with unevenly loaded BSs .................. 78

Figure 6-7: DS throughput with local load of 0.5 ....................................................... 80

Figure 6-8: DS throughput with local load of 1.0 ....................................................... 80

Figure 6-9: US Throughput with no local LAN Traffic .............................................. 81

Figure 6-10: US Throughput with 0.5 local LAN Traffic ............................................ 81

Figure 6-11: Average ETE Delay when one BS sends high data rate traffic to other BS in the .............................. 82

Figure 6-12: Average ETE Delay when one BS sends traffic to other BS in the EPON-based... 82

Figure 6-13: IEEE 802.16e Scanning Procedure ....................................................... 86

Figure 6-14: Non-coordinated Association (Association level 0) ................................. 88

Figure 6-15: The network entry process. Exchange message depend upon IEEE 802.16e standard 92

Figure 6-16: Optimized Distributed Hard Handover (ODHHO) Algorithm ................. 94

Figure 6-17: Handover Delay Comparison: Typical Vs. EPON-Based Distributed Architectures 97

Figure 6-18: Simulation Setup .................................................................................. 98
Figure 6-19: WiMAX Mobility Analysis in Typical Architecture ............................................. 100

Figure 6-20: Initial Ranging activity, Scanning interval activity, BS ID to handover and SNR. 101

Figure 6-21: WiMAX handover activity zoomed in close to frame level ............................. 102

Figure 6-22: Throughput analysis between Typical and Distributed architectures .......... 103

Figure 6-23: Average Throughput analysis between typical and distributed architectures ..... 103

Figure 6-24: MS UL Throughput analysis Typical Vs. Distributed Architectures .............. 104

Figure 6-25: MS DL Throughput analysis Typical Vs. Distributed architectures ............... 104

Figure 6-26: Handover Delay Comparison Typical Vs. Distributed architecture .............. 105

Figure 6-27: Average Handover Analysis Typical Vs. Distributed architectures .............. 106

Figure 6-28: Simulation Setup ........................................................................................ 109

Figure 6-29: MS Traffic Received comparison .................................................................. 110

Figure 6-30: MS Average Traffic Received ....................................................................... 110

Figure 6-31: Simulation Setup ........................................................................................ 111

Figure 6-32: MS’s Traffic Received comparison ............................................................... 113

Figure 6-33: MS’s Average traffic Received ................................................................. 113

Figure 6-34: MS’s Total Average traffic Received ......................................................... 114

Figure 6-35: MS’s Total Average traffic Received ......................................................... 116

Figure 6-36: UL Handover Throughput EPON-based Vs. typical-WiMAX ...................... 116

Figure 6-37: Packet Loss analysis EPON-based Vs. Typical-WiMAX ............................ 117

Figure 6-38: Average Handover delay EPON-based Vs. Typical WiMAX ..................... 117
LIST OF TABLES

Table 2-1: Bandwidth assignments for LTE ....................................................... 22

Table 2-2 WiMAX Parameters ................................................................. 26

Table 3-1: Parameters used in the Model ................................................ 40

Table 6-1: Unevenly Loaded Base Stations Scenarios .............................. 73

Table 6-2: Simulation Parameters for Handoff Scenarios .......................... 107
1 Introduction
1.1 Overview
The growing demand for a myriad of non-voice services including text messaging (SMS), email, multi-media messaging (MMS), Web and WAP access, mobile music, mobile TV, mobile video, and IPTV, have contributed to an explosion in mobile data traffic. With Fourth-Generation (4G) mobile WiMAX and Long-Term Evolution (LTE) rollout accelerating and LTE-Advanced (LTE-A) on the horizon, the phenomenal growth of mobile backhaul capacity required to transport these high-bandwidth applications, with peak air throughput per user reaching 20-500 Mbps, is far beyond what can realistically be achieved using today’s voice-centric mobile backhaul transport networks [1-5]. Mobile backhaul, sometimes referred to as the Radio Access Network (RAN), is utilized to backhaul traffic from individual base stations (BSs) to the Radio Network Controller (RNC), which then connect to the mobile operator’s core network or gateway.

The fundamental problem is that the majority of today’s RANs are built on legacy TDM-based circuit-switched infrastructure, where each BS is connected directly to the RNC via dedicated point-to-point T1/E1 lines. Optimized for slow growing, narrowband, circuit-switched voice traffic, current circuit switched T1/E1 wireline and microwave backhaul technologies can neither meet the capacity requirements of widespread 4G nor cope with the dynamic and highly fluctuant traffic pattern of the emerging data-centric 4G mobile multimedia services.

For instance, the average US carrier’s cellular BS might have something in the neighborhood of 4 T1s feeding them for a total of about 6 Mbps of aggregate bandwidth in and out of the BS. This begs the question of how these most-widely talked about “100-400 Mbps mobile connections per subscriber in the downstream and upstream directions, respectively” can realistically be supported when the maximum connection beyond the current BS is a mere 6Mbps? This will certainly require rapid migration from today’s legacy circuit switched T1/E1 wireline and
microwave backhaul technologies to a new fiber-supported, all-packet-based mobile backhaul infrastructure.

In addition, RF resources (RF transceivers and dedicated timeslots assigned to these across the T1/E1 links) are designed for worst-case traffic scenario (peak hour’s traffic), so most of the time these dedicated resources are idle and wasted. There is no way to dynamically move capacity from a heavily loaded BS to a lightly loaded/idle BS even within the same cell site. In other words, additional backhaul expense results from unused bandwidth that is often stranded in the “wrong place at the wrong time”. Hence, to cope with the growing and dynamic nature of the emerging mobile data-centric traffic and services, it is essential that future RAN architectures must support dynamic bandwidth allocation (DBA) and sharing in both downstream and upstream directions.

Exacerbating the problem is that, in contrast to the typically centralized 2G/3G RAN architecture, the 4G standards require new distributed RAN architectures and further create a requirement to fully meshing the LTE BSs [1-3]. Thus, the more distributed architecture associated with 4G necessitates fundamentally different RAN design requirements. Specifically, the efficient applicability of many of the 4G critical features including rapid handoff and Coordinated Multipoint (CoMP) transmission and reception techniques between neighbors BSs depends to a great extent on the backhaul characteristics (latency and capacity), which is driven by the transport technology (e.g., optical fiber, microwave or copper-based technologies) and the RAN topology.

Clearly, a cost effective fiber supported all-packet-based mobile backhaul RAN architecture that is compatible with these inherently distributed 4G RAN architectures, that is the focus of this
thesis, is needed to efficiently scale current mobile backhaul networks. This requires rapid migration from today’s legacy circuit switched T1/E1 wireline and microwave backhaul technologies to a new fiber-based mobile backhaul infrastructure. However, deploying a green fiber-based mobile backhaul infrastructure is a costly proposition mainly due to the significant cost associated with digging the trenches in which the fiber is to be laid. These, along with the inevitable trend towards all-IP/Ethernet transport protocols and packet switched networks, have prompted many carriers around the world to consider the potential of utilizing the existing fiber-based Passive Optical Network (PON) access infrastructure as an all-packet-based converged fixed-mobile optical access networking transport architecture to backhaul both mobile and typical wireline traffic.

Passive Optical Network (PON)-based fiber-to-the-curb/home (FTTC/FTTH) access networks are being deployed around the globe based on two Time-Division Multiplexed (TDM) standards: ITU G.984 Gigabit PON (GPON) and IEEE 802.ah Ethernet PON (EPON) [6-11]. A PON connects a group of Optical Network Units (ONUs) located at the subscriber premises to an Optical Line Terminal (OLT) located at the service provider’s facility. ITU GPON standard supports asymmetric 2.5 Gbps downstream (DS)/1.25 Gbps upstream (US) channel capacity, while IEEE EPON standard supports symmetric 1 Gbps. Since all users attached to a TDM-PON share a single DS/US transmission channel, the average dedicated bandwidth that can be supported in either direction is usually limited to a few tens of Mbps per user. While current generation PON technologies seem to offer a satisfactory solution for residential bandwidth demands, they certainly will not be able to meet future demands as well as the emerging business and mobile backhaul services.
To address this issue, the IEEE and ITU Full Service Access Network (FSAN) community have commissioned studies on defining possible smooth migration scenarios from the current Gigabit-class PON systems toward the next-generation of PON (NG-PON) access systems that are compatible with the current PON systems but with much higher bandwidth [12-16]. The outcome of these studies, which is eloquently summarized in [12-15], endorses two potential candidate system architectures for NG-PON. The first is a 10G TDM PON evolutionary growth architecture, termed NG-PON1, which supports coexistence with legacy PONs on the same optical distribution network (ODN) and is viewed as a mid-term upgrade. The second is a revolutionary disruptive architecture (termed NG-PON2) with no requirements in terms of coexistence with current PON on the same ODN, and is regarded as a longer-term solution.

NG-PON1 is further subdivided into two different 10G standards: the ITU XG-PON (X taken as the roman sign for 10) and the IEEE 10G EPON. The IEEE 10G EPON supports both asymmetrical 10/1 and symmetrical 10/10 Gbps solutions. Because the 10 Gbps US burst transmission is more challenging than either 1 or 2.5 Gbps, both the IEEE and ITU symmetrical 10/10 Gbps solutions are expected to be practical in a longer time frame and it will also be more expensive.

1.2 Thesis Motivation

Given the large investments many fixed-line carriers are making or have already made in PON-based FTTH/FTTC access infrastructure, the combination of PON and native Ethernet, albeit with carrier-class enhancements, with a fiber-based access infrastructure is the most promising mix of technologies that ensures a cost-effective and future-proof converged fixed-mobile access transport infrastructure. The economic advantage of utilizing the existing fiber-based PON
access infrastructure with Ethernet functionality is quite compelling compared to the choice of continuing to invest in legacy technology and/or the costly proposition of blinding up a new packet-based mobile backhaul infrastructure.

Numerous research efforts have recently emerged to address the integration of PON and wireless broadband access technologies via proposing numerous hybrid Fiber-Wireless (FiWi) network architectures that will enable the support of fixed-mobile applications and services independent of the access infrastructure [17-22]. These architectures have typically assumed utilizing the fiber-based PON access infrastructure to backhaul mobile traffic, in addition to typical wireline traffic. Most of these architectures, however, have typically assumed a centralized RAN architecture for the wireless segment of the hybrid architecture (e.g. 3G, WiFi). Since both the wireless and wireline segments of the hybrid architecture are centralized, the two access technologies are typically operated independently by considering the BS a generic user attached to an ONU and/or or collocated with it.

Interconnecting the BSs and ONUs via a common standard interface (e.g., Ethernet) is the main additional component required to implement such hybrid converged fixed-mobile access architecture from the two independent access systems. Thus, the important problem of how to efficiently interwork between the two access technologies to optimize the overall hybrid system performance is totally disregarded. In this case, the main and only performance advantage gained from such hybrid architecture is the typical utilization of the fiber-based PON access infrastructure merely as a dumb high bandwidth pipe to backhaul mobile traffic, in addition to typical wireline traffic.
However, if the RAN architecture of the wireless segment is fully distributed (e.g., 4G LTE), as this work will show, a framework to efficiently interwork the two access technologies is now indispensable to enable the typically centralized PON architecture to support the distributed 4G radio network control and management (NCM) operations as well as to fully mesh the 4G LTE BSs (i.e., to conform with the 4G LTE standards). As this work will show, devising a truly unified NG-PON-4G access networking architecture that conforms to both the typically centralized PON and the emerging distributed 4G access standards significantly enhances the overall hybrid system performance but certainly adds a number of challenging dimensions to the problem.

While the economics for commercially deploying TDM-based NG-PONs in the access arena as a near-term 4G mobile backhaul RAN infrastructure are quiet compelling, however, several key outstanding technical hurdles must be addressed first before mainstream TDM-PONs evolve as viable hybrid fixed-mobile access networking technology that enables the support of a truly unified PON-4G access networking architecture. These include:

1. TDM-PON is a centralized access architecture—relying on a component at the distant OLT to arbitrate upstream traffic, while 4G is a distributed architecture where, in particular, the 4G LTE/LTE-A standard requires a new distributed RAN architecture and further create a requirement to fully meshing the BSs (the X2 interface for LTE BS-BS handoffs requires a more meshed architecture) [1-3]. Thus, a PON-based 4G mobile backhaul RAN must be capable of supporting a distributed architecture as well as distributed NCM operations. Exacerbating the problem is that mainstream PONs are typically deployed as tree topologies. However, tree-based topology can neither supports distributed access architecture nor intercommunication among the access nodes (ONUs)
attached to the PON. The key challenge in devising a PON-based 4G mobile backhaul RAN architecture or a truly unified PON-4G access networking architecture is how to efficiently interwork and reconcile the traditionally centralized PON’s architecture and NCM operations with the typically distributed 4G’s architecture and NCM operations.

2. A typical tree-based PON topology can’t cope with LTE-A CoMP technology requirements on backhaul links. This is because the LTE BSs are interconnected via the standardized X2 interface, which is carried over the mobile backhaul (PON-based RAN). To efficiently implement CoMP features, specific control and signaling among the master BS and the rest of the cluster BSs (or neighbors BS) will be required to exchange channel state information (CSI) and commands with very low latency (X2 delay should be in a range of 1 ms or lower [5]). This stringent requirement on the X2 latency can’t be adequately met using the typical tree-based PON topology; a more meshed topology is rather required.

3. Due to the inherent lack of simple and efficient resilience capabilities in tree-based PON topologies, specifically against failures in the distribution network, service resilience over previous generations of PONs has not been a strong requirement from operators. Since a single wavelength failure may affect the premium services delivered to thousands of fixed-mobile end-users, the reliability offered by a converged access network to the services and customers it supports is one of the most important considerations in designing and deploying such a converged transport network. Thus, typical tree-based PON topologies are not suitable for truly resilient converged access architecture.
1.3 Thesis Statement

This thesis examines the technological requirements and assesses the performance analysis and feasibility for deploying TDM-based NG-PON solutions in the access arena as multiservice, all packet-based 4G mobile backhaul RAN and/or converged fixed-mobile optical networking architecture. Specifically, this work proposes and devises a simple and cost-effective 10G-EPON-based 4G mobile backhaul RAN architecture that efficiently transports and supports a wide range of existing and emerging fixed-mobile advanced multimedia applications and services along with the diverse quality of service (QoS), rate, and reliability requirements set by these services. The techno-economics merits of utilizing PON-based 4G RAN architecture versus that of traditional 4G (mobile WiMAX and LTE) RAN will be thoroughly examine and quantified. The proposed architecture builds upon a novel fully distributed approach to the design and implementation of NG-PONs, and so is likely to be disruptive.

To achieve our objective, we utilize the existing fiber-based PON access infrastructure with novel ring-based distribution access network and wireless-enabled OLT and ONU as the multiservice packet-based 4G mobile backhaul RAN infrastructure. Specifically, to simplify the implementation of such a complex undertaking, this work is divided into two sequential phases. In the first phase, we examine and quantify the overall performance of the standalone ring-based 10G-EPON architecture (just the wireline part without overlaying/incorporating the wireless part (4G RAN)) via modeling and simulations. We then assemble the basic building blocks, components, and sub-systems required to build up a proof-of-concept prototype testbed for the standalone ring-based EPON architecture. The testbed will be used to verify and demonstrate the
performance of the standalone architecture, specifically, in terms of power budget, scalability, and reach.

In the second phase, we develop an integrated framework for the efficient interworking between the two wireline PON and 4G mobile access technologies, particularly, in terms of unified NCM operations. Specifically, we address the key technical challenges associated with tailoring a typically centralized PON-based access architecture to interwork with and support a distributed 4G RAN architecture and associated radio NCM operations. This is achieved via introducing and developing several salient-networking innovations that collectively enable the standalone EPON architecture to support a fully distributed 4G mobile backhaul RAN and/or a truly unified NG-PON-4G access networking architecture. These include a fully distributed control plane that enables intercommunication among the access nodes (ONUs/BSs) as well as signaling, scheduling algorithms, and handoff procedures that operate in a distributed manner.

It is important to emphasize that the proposed architecture eloquently complies with both of LTE and LTE-A’s standards via the purposely-selected simple ring topology (versus a typical tree-based PON topology) for the distribution network, which is the key for addressing the key aforementioned hurdles. This is because: 1) it enables direct intercommunication/connectivity among the access nodes (ONUs/eNBs), allowing for the support of a distributed PON-4G RAN access architecture as well as for simply fully meshing the access nodes, in conformity with the 4G LTE standards; 2) it minimizes the X2 interface latency allowing for harnessing the highest CoMP gains; and 3) the inherent self-healing mechanism of a ring architecture provides the critically needed network resiliency feature that guarantees the reliable delivery of both fixed and mobile services.
The ring-based local access network topology along with the developed networking innovations including the fully distributed control plane and NCM operations collectively enable the proposed EPON-based 4G RAN architecture to support several novel salient networking features including:

1. The overall performance of proposed EPON-based 4G backhaul including both the RAN and MPC (Evolved Packet Core (EPC) per 3GPP LTE’s standard) is significantly augmented compared to that of the typical 4G, specifically, in terms of handoff capability, signaling overhead, overall network throughput and latency, and QoS support.

2. It enables redistributing some of the intelligence and NCM operations currently centralized in the Mobile Packet Core (MPC) platform out into the access nodes of the mobile RAN. Specifically, as this work will show, it enables offloading sizable fraction of the mobile signaling as well as actual local upstream traffic transport and processing (LTE bearers switch/set-up, retain, and tear-down and associated signaling commands from the BSs to the EPC and vice-versa) from the EPC to the access nodes (ONUs/BSs). This has a significant impact on the performance of the EPC. First, it frees up a sizable fraction of the badly needed network resources as well as processing on the overloaded centralized serving nodes (AGW) in the MPC. Second, it frees up capacity and sessions on the typically congested mobile backhaul from the BSs to the EPC and vice-versa.

3. It completely eliminates the typical utilization of the 10 Gbps upstream burst-mode transmitter/receiver and associated design challenges at the ONU/OLT. This facilitates and expedites the near-term deployment of symmetrical 10/10 Gbps NG-PON solutions. It also alleviates the typical limited US power budget problem, specifically for the most stringent 10G-
EPON high power budget (> 30 dB) class specifications (PR/PRX30) for symmetric DS/US 10 Gbps transmission.

4. The proposed TDM-based EPON transport infrastructure is transparent to the type of network traffic utilized (wireline, mobile, or a mix of both). Thus, the significance of the proposed PON-based networking architecture is that it can be used as a generic access transport infrastructure to backhaul either only wireline traffic (typical TDM/WDM-PONs), mobile traffic (mobile backhaul RANs), and/or a mix of both (converged fixed-mobile architecture).

Overall, the proposed NG-PON architecture constitutes a complete networking paradigm shift from the typically centralized PON’s architecture and OLT-based NCM operations to a new disruptive fully distributed PON’s architecture and NCM operations in which all the typically centralized OLT-based PON’s NCM operations are migrated to and independently implemented by the access nodes (ONUs) in a distributed manner. This requires migrating most of the typically centralized wireline and radio control and user-plane functionalities such as DBA, queue management and packet scheduling, handover control, radio resource management, admission control, etc., typically implemented in today’s OLT/RNC, to the access nodes (ONUs/4G BSs). Thus, under the proposed fully distributed scenario, each ONU/4G BS independently implement typical wireline/radio control and user-plane functionalities in a distributed approach without resorting to a central control entity (OLT/LTE’s access gateway (AGW), in conformity with 4G LTE standards).

Most of the results presented in this thesis are obtained through extensive computer simulations and modeling using substantial system simulator development, which is carried out during the course of this work. This includes both mathematical modeling considerations as well as
software design, implementation, and testing. The system parameters used here to assess the overall performance of the proposed architecture are identical to the performance evaluation parameters specified in Mobile WiMAX system evaluation documents and WiMAX profiles.

Finally, in this work, though we have chosen Ethernet-based PON and mobile WiMAX as representative techniques for fixed PON and 4G mobile access technologies, the proposed architecture and related operation principles are also applicable to other PON and 4G access networks such as GPON and LTE. In fact, the proposed architecture is specifically tailored to support the distributed nature of LTE-A RAN architecture and to harness the highest CoMP gains.

1.4 Thesis Organization

Organizational outline of the thesis is as follows:

Chapter 2 starts with the introduction of 4G Mobile Wireless Technologies LTE followed by WiMAX. At the end both technologies are compared and future trends are investigated.

Chapter 3 starts with the discussion of major problems associated with a centralized architecture as the “single-point of failure” problem i.e., the failure of the OLT software will bring down the whole access network. Another major problem discussed is that the PON architecture is typically centralized but 4G RAN architecture is intrinsically distributed. We propose distributed solutions to this problem, and in the process to prove that these distributed networking architectures and the associated bandwidth allocation algorithms and protocols have characteristics that make them far better suited for provisioning Quality of Service (QoS) schemes necessary for properly handling data, voice, video, and other real-time streaming advanced multimedia services over a single line.
Chapter 4 discusses EPON-Based 4G Mobile Backhaul RAN Architecture. Moreover it also illustrates some of the key technical challenges associated with devising a truly unified fixed-mobile NG-PON-WiMAX/LTE/ (A) access transport architecture that is built on top of a typically centralized PON infrastructure.

Chapter 5 introduces the key networking building blocks that enable the standalone ring-based EPON architecture to support a fully distributed 4G mobile backhaul RAN architecture and/or converged/unified fixed-mobile optical access networking infrastructure.

Chapter 6 describes Key Technological and Economic Advantages Enabled By the Distributed EPON-Based RAN Architecture.

The distributed ring-based architecture along with the supporting control plane enable the proposed EPON-based RAN architecture to support several salient networking features that collectively significantly augment the performance of both the RAN and MPC in terms of handoff capability, overall network throughput, stability, and latency, and QoS support. capabilities. This chapter illustrates these advantages with extensive simulation results.
Chapter 2

2 Overview of 4G Wireless Technologies-WiMAX & LTE
2.1 Introduction
Mobile networks and services have gone beyond voice-only communication services and are rapidly evolving towards data-centric services. Emerging mobile data services are expected to see the same explosive growth in demand that Internet and wireless voice services have seen in recent years. Within few years, global broadband wireless access services and connections will surpass those of DSL, cable modem, and fiber combined. To support such a rapid increase in traffic, active users, and advanced multimedia services implied by this growth rate along with the diverse quality of service (QoS) and rate requirements set by these services, mobile operator need to rapidly transition to a simple and cost-effective, flat, all IP-network.

This has accelerated the development and deployment of new wireless broadband access technologies including High-Speed Packet Access (HSPA) and fourth-generation (next generation) mobile WiMAX and cellular Long-Term Evolution (LTE). The emerging next generation mobile access technology is the core of the envisioned flat, all-IP network architecture that is capable of delivering speeds comparable to or better than current fixed-line broadband access systems – up to 15-100 Mb/s peak air throughput per user. The future of next generation mobile broadband is a tale of two technologies: LTE and mobile WiMAX (Worldwide Interoperability for Microwave Access).

This chapter starts with the introduction of LTE followed by WiMAX. At the end both technologies are compared and future trends are investigated.
2.2 LONG TERM EVOLUTION (LTE)

2.2.1 LTE Basics

The trend of ever increasing transmission bandwidths is challenging the limits of current 3G networks, hence it was decided by 3GPP (3rd generation partnership program) standardization body in 2005 to start work on next generation wireless network design that is only based on packet-switched data transmission. LTE is the latest standard in the mobile network technology tree that is being implemented within the Third Generation Partnership Project (3GPP) to ensure the competitiveness of 3G for the next 10 years and beyond. LTE supports both time-division duplex (TDD) and frequency division duplex (FDD). Moreover it supports a flexible and scalable bandwidth e.g., 1.25, 5, 10 and 20 MHz. Moreover LTE has a very flexible radio interface [1, 2].

In LTE base station is referred to as enhanced NodeB (eNodeB) in order to differentiate it from UMTS (Universal mobile telecommunication system) base station which is known as NodeB. Enhanced NodeB (eNodeB) base stations are made more intelligent than NodeB by removing Radio Network Controller (RNC) and transferring the functionality to eNodeB and partly to the core network gateway. In LTE the base stations can also perform handovers as they can communicate directly over X2 interface. S1 interface connects eNodeB to the gateway nodes i.e., between radio network and core network. It is completely based on IP protocol. The gateway between radio access network and core network is divided into two entities Serving Gateway (Serving-GW) and the Mobility Management Entity (MME). MME is the control plane (c-plane) entity is mainly responsible for subscriber mobility, session management signaling, location tracking of mobile devices and selection of a gateway to the internet when mobile requests IP
address from the network. On the other hand Serving-GW is responsible for user plane (u-plane). Both components can be implemented on the same hardware or separated. If implemented separately, S11 interface is used to communicate between them. Basic LTE network architecture [3] is shown in following figure 2-1.

MME: Mobility Management Entity
S-GW: Serving Gateway
P-GW: Packet Data Network Gateway
HSS: Home Subscriber Server
SCP: Service Control Point
AGW: Access Gateway
eNB: Enhanced NodeB
UE: User Equipment

**Evolved packet core (EPC)**

*Figure 2-1: LTE Network Architecture*
S6 interface is between MME and database that stores subscription information, referred to as Home Subscriber Server (HSS). In LTE, the router at the edge of the wireless core network is known as Packet Data Network Gateway (PDN-GW) and the interface between PDN-GW and MME / Serving-GW is called S5. It uses GTP-U (user) protocol to tunnel user data from / to the Serving-GWs and the GTP-S (Signaling) protocol for the initial establishment of a user data tunnel and subsequent tunnel modifications when the user moves between cells that are managed by different Serving-GWs.

For air interface, LTE uses Orthogonal Frequency Division Multiple Access (OFDMA) in the downlink data transmission. In OFDMA, a big data stream is transmitted by using many narrow band sub-carriers simultaneously. The sub-carriers are spaced apart at fixed frequencies (15 KHz). This spacing provides orthogonality among carriers, as shown in Figure 2-2. Because many bits of data are transmitted in parallel, the transmission speed of each sub-carrier can be much lower than the overall data rate. This not only minimizes the multipath fading but also the effect of multipath fading and delay spread become independent of the channel bandwidth used. This is because the bandwidth of each sub-carrier remains same and only the number of sub-carriers is changed for different achievable overall bandwidth. Moreover OFDMA has more advantages like high spectral efficiency. The most common modulation techniques used are binary phase-shift keying (BPSK), quadrature phase-shift keying (QPSK) and quadrature amplitude modulation (QAM).
For OFDMA downlink transmission, a mathematical function Inverse Fast Fourier Transform (IFFT) transforms the signal from frequency domain to time domain. The resulting signal is then modulated and amplified and transmitted in the air. When the signal is received by the receiver, it first demodulates and amplifies the signal. After this the signal is converted back from time domain to frequency domain using Fast Fourier Transform (FFT). The multiple access (MA) in OFDMA refers to the data that is sent in the downlink is received by several users simultaneously. This is accomplished by the use of control messages to inform mobile devices, waiting for data, which part of data is addressed to them and which part they can ignore. On the physical layer it means the use of modulation schemes ranging from QPSK over 16QAM to 64QAM can be quickly changed for different sub-carriers to fulfill different reception conditions.

In LTE, for uplink transmission, a different transmission scheme is used as compared to in the downlink. This is known as Single Carrier Frequency Division Multiple Access (SC-FDMA). This is because OFDMA inherently suffers from high peak to average power ratio (PARP) which can drain the mobile device battery quickly. Since mobile device should consume as little energy as possible, a different transmission technique SC-FDMA is proposed for the uplink transmission. In general this scheme is similar to OFDMA but has much lower PARP. This is the
reason SC-FDMA is selected for uplink transmission. SC-FDMA also transmits data over the air interface in many subcarriers, but adds an additional processing step. A number of input bits are grouped and then passed through FFT first and then output of FFT is fed into IFFT block. Since not all the subcarriers are used by the mobile station, many of them set to zero. On the receiver side the signal is amplified, demodulated and then fed into FFT block. The resulting signal is fed into IFFT block to counter the effect of additional step in the transmission. The resulting time domain signal is fed into detector block which recreates the original signal bits.

2.2.2 Physical Parameters and Frame Structure

For LTE, physical parameters are chosen as follows:

- OFDM symbol duration, 66.667 µs
- Subcarrier spacing, 15 kHz
- Standard cyclic prefix (CP), 4.7 µs
- Extended cyclic prefix (CP), 16.67 µs

The cyclic prefix (CP) is transmitted before each OFDM symbol to prevent inter-symbol interference (ISI) which is evident because of different transmission paths of varying lengths.

Moreover in LTE different channel bandwidths ranging from 1.25 MHz to 20 MHz Table 2-1 shows the standardized transmission bandwidths, the number of subcarriers used and the FFT size used and physical Resource Block (PRB) for each bandwidth. Physical Resource Block (PRB) is the smallest element of resource allocation assigned by the base station scheduler.
Table 2-1: Bandwidth assignments for LTE

<table>
<thead>
<tr>
<th>Bandwidth (MHz)</th>
<th>Number of subcarriers</th>
<th>Subcarrier Bandwidth (kHz)</th>
<th>FFT size</th>
<th>Physical Resource Block (PRB)</th>
<th>PRB Bandwidth (kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25</td>
<td>76</td>
<td></td>
<td>128</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>151</td>
<td></td>
<td>256</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>301</td>
<td>15</td>
<td>512</td>
<td>25</td>
<td>180</td>
</tr>
<tr>
<td>10</td>
<td>601</td>
<td></td>
<td>1024</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>901</td>
<td></td>
<td>1536</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1201</td>
<td></td>
<td>2048</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

LTE generic frame structure is shown in figure 2-3. It is evident from the figure that LTE frame duration is 10 ms. It is then divided into 10 sub frames of 1 ms duration each. Each sub frame is further subdivided into two slots of 0.5 ms each. Each Slot of 0.5 ms consists of 12 subcarriers and 6 or 7 OFDM symbols depending upon either standard or extended cyclic prefix (CP) is used. When extended cyclic prefix is used then the number of OFDMA symbols reduced to 6. The grouping of 12 subcarriers results in PRB bandwidth of 180 kHz.

Two slots that grouped together to form a sub frame which is also known as Transmit Time Interval (TTI). In case of Time Division Duplex (TDD) operation, the sub frame can be used for downlink or uplink. This is decided by the network which frames are used for downlink or uplink. However in LTE most networks likely to use Frequency Division Duplex (FDD) in which separate bands are used for uplink and downlink.
Data is mapped to subcarriers and symbols, which are arranged in the time and frequency domain in a resource grid. LTE physical resource block is shown in Figure 2-4.
2.2.3 MIMO Transmissions

LTE standard defines the use of MIMO technology. Transmission of several independent signals over the same frequency band is also referred to as MIMO or multiple input multiple output. 

LTE standard defines two and four transmissions over the same band, which needs 2 or 4 antennas at both receiver and transmitter side respectively. A comprehensive mathematical treatment of MIMO is given in 4. These transmissions are known as 2 x 2 MIMO and 4 x 4 MIMO. Since MIMO channels are separated from each other, 2 x 2 MIMO can increase overall data rate by two and likewise 4 x 4 MIMO by four times. However LTE is only used in the downlink transmissions since for uplink transmissions it is difficult to use MIMO for mobile devices because of limited antenna size and power constraints.

2.3 WORLDWIDE INTEROPERABILITY FOR MICROWAVE ACCESS (WIMAX)

Like LTE, WiMAX is another successor of 3.5G wireless network technologies and is based on IEEE 802.16 air interface. While LTE is the evolution of 3G wireless networks, WiMAX is a new standard which is not backward compatible and is suitable for new operators. The WiMAX network architecture is fully based on IP and greatly simplifies the network architecture design and deployment.

2.3.1 WiMAX Network Architecture

When WiMAX network started to deploy, it targeted only stationary devices with roof mounted antennas or indoor WiMAX routers. That version of WiMAX is known as IEEE 802.16-2004 or 802.16d. Since these networks designed to handle stationary devices, so no handovers of connections between base stations are required.
Second version of WiMAX known as IEEE 802.16-2005 or 802.16e or mobile WiMAX is the advanced air interface which not only supports mobility of subscribers but also handovers between base stations. Since mobile subscribers and handovers require more administration so it became necessary to define the network beyond base stations. As IEEE is only responsible for the air interface so that task was carried over by WiMAX Forum 5. Moreover in order to ensure interoperability of devices and components among variety of vendors, WiMAX forum also conducts certification program for base stations, network equipments and end user devices.

Figure 2-5 shows one of the practical WiMAX network architecture [3].
As evident from the figure 2-5, R1 reference point is the air interface between mobile devices and base station. This is fully based on IP protocol. The protocol used over this interface is either IEEE 802.16d for only fixed clients or IEEE 802.16e for both mobile and fixed clients. Base stations can carry out smooth handovers among them using R8 interface. R6 interface is between base stations and Access Service Network Gateway (ASN-GW). ASN-GW is the gateway between radio network and core network. In LTE, this is comparable to Access Gateway (AGW). R3 interface is defined between ASN-GW and core network.

2.3.2 IEEE 802.16-2005 Air Interface

Some of the parameters of IEEE 802.16-2005 which is also known as mobile WiMAX are given in the Table 2-2. Complete description of IEEE 802.16-2005 can be found in reference6. As LTE, WiMAX employs OFDMA in the downlink but for the uplink it differs from the LTE as it defines to use OFDMA also in the uplink. LTE uses SC-FDMA in the uplink. The mobile WiMAX standard supports both TDD and FDD modes but initial deployments are only using TDD.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Bandwidth (MHz)</td>
<td>1.25</td>
</tr>
<tr>
<td>Number of subcarriers</td>
<td>85 421</td>
</tr>
<tr>
<td>FFT Size</td>
<td>128 512</td>
</tr>
<tr>
<td>Subcarrier spacing (kHz)</td>
<td>10.94</td>
</tr>
<tr>
<td>OFDM symbol duration (μs)</td>
<td>91.4</td>
</tr>
<tr>
<td>Cyclic Prefix (μs)</td>
<td>11.4</td>
</tr>
</tbody>
</table>
OFDMA TDD frame structure is shown in the following figure 2-6. In OFDMA data transmissions to users are multiplexed in both time and frequency because of number of available sub channels.

![OFDMA TDD Frame Structure](image)

Figure 2-6: OFDMA TDD Frame Structure

At the beginning of the frame, Downlink-MAP (DL-MAP) informs devices when and where the data is scheduled for them in the frame. Uplink-MAP (UL-MAP) in the frame assigns uplink transmissions opportunities for this and the frames to follow.

### 2.3.3 MIMO Transmissions

Like LTE, WiMAX standard also supports the use of MIMO transmissions e.g., 2 x 2 MIMO to increase the overall data rate by two. Moreover MIMO is only implemented in the downlink direction by the base stations. In the uplink direction MIMO is not implemented due to limited antenna size and power constraints.
2.4 COMPARISON OF NEXT GENERATION MOBILE WIRELESS TECHNOLOGIES, LTE AND WIMAX

LTE and WiMAX perform in a very similar manner as both use OFDM modulation and also have similar radio parameters. The major difference between two is that WiMAX will be deployed in TDD mode first while LTE will be using FDD mode because of spectrum availability and also because LTE is basically evolution of 3G technology. Another difference is that LTE has different uplink scheme, namely SC-FDMA, making it more power efficient for mobile devices. However it is still to be seen that how much difference it will make. On the other hand WiMAX has five QOS (Quality of Service) parameters 7, UGS (Unsolicited Grant Service), ERTP (Extended Real Time Polling Service), RTP (Real Time Polling Service), NRTP (Non Real Time Polling Service) and BE (Best Effort). In comparison of LTE QOS parameters of two GBR (Guaranteed Bit Rate) and NGBR (Non-Guaranteed Bit Rate), which are further subdivided into small types which are defined by their priority.

Mobile WiMAX and LTE are two different (but not necessarily competing) technologies that will ultimately be used to achieve data speeds of up to 100 Mbps. Speeds that are fast enough to potentially replace wired broadband connections with wireless, and enable services such as HDTV on mobiles and TVs without the need for a fixed-line or dish in the home, as well as a host of other exciting services currently seen as too bandwidth-hungry to be delivered using existing mobile technologies. They are both in different stages of development. WiMAX is widely recognized as being the first that will be brought to market. The world's first large scale mobile WiMAX deployment has already started to roll in the US in spring 2008 (Sprint and ClearWire). However, the crucial difference is that, unlike WiMAX, which requires a new
network to be built, LTE runs on an evolution of the existing UMTS infrastructure already used by over 80 per cent of mobile subscribers globally. This means that even though development and deployment of the LTE standard may lag Mobile WiMAX, it has a crucial incumbent advantage.

One of the biggest obstacles to widespread WiMAX deployments is the lack of available high quality spectrum. In the US, Sprint benefits greatly from its 2.5 GHz spectrum holdings. This relatively low-frequency band allows greater coverage per base station since signals travel much further than at higher frequencies. This result in fewer base stations needed, making WiMAX cheaper to deploy in the US than in other markets that don't have access to the same spectrum. Even given the availability of 2.5 GHz spectrum, for Sprint's network to provide nationwide coverage it will require more than 60,000 base stations across the US.

In Europe, bandwidth below 2.5GHz is scarce and mostly occupied by analogue TV and current GSM mobile signals. Therefore, until now most European WiMAX trials and licenses have been limited to the 3.5 GHz or even 5 GHz bands with often disappointing results, which is why we haven't seen anywhere near as much WiMAX traction in Europe as the US. It may not be until after analogue broadcast signals are switched off across Europe (with the UK scheduled for 2012) that sub 2.5 GHz spectrum becomes available and we start to see large-scale European WiMAX deployments.

So which technology will ultimately prevail? Is the battle for new mobile broadband networks will be a contest between WiMAX and LTE? The way the battle now is shaping up; WiMAX has the speed to market advantage. But LTE is embraced by most of the global mobile carrier industry. It is also debatable that LTE is more ‘risk-free’ than WiMAX because it will run on an
evolution of existing mobile infrastructure. Also, mobile operators will be able to use their experience from current 3G and HSDPA networks to carry out the incremental fine-tuning necessary to ensure that the rollout of LTE will deliver on user expectations. Also in Europe it has the advantage of being unaffected by the lack of available spectrum. However, the recognition of WiMAX as an IMT-2000 technology by the ITU in October 2007 is a significant step, that in the future may help WiMAX to gain a foothold in today's UMTS spectrum and so close the spectrum availability gap, but the full impact of this move has yet to unfold.
3 Fully Distributed Ring-Based 10G-EPON

Architecture
3.1 Introduction

To date, mainstream Ethernet Passive Optical Network (EPON) bandwidth allocation schemes as well as the new IEEE 802.3ah Ethernet in the First Mile (EFM) Task Force specifications have been centralized, relying on a component in the central office (Optical Line Termination (OLT)) to provision upstream traffic. Hence, the OLT is the only device that can arbitrate time-division access to the shared channel. Since the OLT has global knowledge of the state of the entire network, this is a centralized control plane in which the OLT has centralized intelligence. One of the major problems associated with a centralized architecture is the “single-point of failure” problem that is the failure of the OLT software will bring down the whole access network. Another major problem is that the PON architecture is typically centralized but 4G RAN architecture is intrinsically distributed. Thus the PON architecture must support a distributed architecture as well as distributed radio network control and management (NCM) operations.

In this section we propose distributed solutions to this problem, and in the process to prove that these distributed networking architectures solutions and the associated bandwidth allocation algorithms and protocols have characteristics that make them far better suited for provisioning Quality of Service (QoS) schemes necessary for properly handling data, voice, video, and other real-time streaming advanced multimedia services over a single line.

3.2 Standalone Ring-Based 10-G-EPON Architecture

3.2.1 Normal State Operation

The standalone architecture refers here to just the wire line segment of the hybrid architecture (10G-EPON) without incorporating the wireless segment (4G LTE RAN). Fig. 3-1 illustrates the standalone ring-based EPON architecture [6]. An OLT is connected to N ONUs via a 20 km trunk feeder fiber, a passive 3-port optical circulator, and a short distribution fiber ring.
the same local access area as in the typical tree-based architecture, the small ring at the end of
the trunk is assumed to have a 1-2 km diameter. The ONUs are joined by point-to-point links in a
closed loop around the access ring. The links are unidirectional: both downstream (DS) and
upstream (US) signals (combined signal) are transmitted in one direction only. The US signal is
transmitted sequentially, bit by bit, around the ring from one node to the next where it is
terminated, processed, regenerated, and retransmitted at each node (ONU). Since US
transmission is based on a TDMA scheme, inter-ONU traffic (LAN data and control messages)
is transmitted along with upstream traffic destined to the OLT (MAN/WAN data) within the
same pre-assigned time slot. Thus, in addition to the conventional transceiver maintained at each
ONU (a λ_up US transmitter (Tx) and a λ_d DS receiver), this approach requires an extra receiver
(Rx) tuned at λ_up to process the received US/LAN signal.

DS signal is coupled to the ring at port 2 of the optical circulator. After recombining it with the
re-circulated US signal via the 2x1 CWDM combiner placed on the ring directly after the optical
circulator, the combined signal then circulates around the ring (ONU_1 through ONU_N) in a Drop-
and-Go fashion, where the DS signal is finally terminated at the last ONU. The US signal
emerging from the last ONU is split into two replicas via the 20:80 1x2 passive splitter (Fig. 3-1)
placed on the ring directly after the last ONU. The first replica (80 %) is directed towards the
OLT via circulator ports 1 and 3, where it is then received and processed by the US Rx (housed
at the OLT), which accepts only MAN/WAN traffic, discards LAN traffic, and process the
control messages, while the second replica (20 %) is allowed to recirculate around the ring after
recombining with the DS signal via the 2x1 CWDM combiner.

The detailed ONU architecture is shown in Fig. 3-2. Each ONU attaches to the ring via the input
port of a 1x2 CWDM DMUX housed at each ONU (incoming signal at point A in Fig.3-2) and
can transmit data onto the ring through the output port of a 2x1 CWDM combiner (outgoing signal at point E in Fig. 3-2). At each ONU, the incoming combined signal is first separated into its two constituent: DS and US signals via the 1x2 CWDM DMUX housed at the ONU. As can be seen from Figure 3-2, the separated US signal is then received and processed via the US Rx housed at the ONU, where it is regenerated and retransmitted along with the ONU’s own local control and data traffic. Note that DS signal is terminated at the last ONU via removing the 2x1 CWDM combiner and 1x2 passive splitter.

As can also be seen from Figure 3-2, the separated DS signal is coupled to the input port of the (10: 90) 1x2 passive splitter, which splits the DS signal into a small (10%) “Drop-signal-portion” and a large (90%) “Express-signal-portion”. The small portion (Drop-Signal) is then received and processed by the DS Rx housed at the ONU. The remaining large portion emerging from the 90% output splitter’s port (Express-Signal) is further transmitted through the ring to the next ONU, where it is, once again, partially split and detected at the corresponding DS Rx and partially transmitted towards the rest of the ring. Note that the Express-Signal recombines again with the retransmitted US signal (all previous ONU’s regenerated US signals plus its own US signal) via the 2x1 CWDM combiner to form the outgoing combined signal (incoming signal for next ONU) that circulates around the ring.

Since the ring is a closed loop, US traffic will circulate indefinitely unless removed. The process of removing, regenerating, and retransmitting the second replica of the US signal at each node (ONU) is implemented as follows: first, the US Rx (housed at each ONU) terminates all US traffic, examines the destination MAC address of each detected Ethernet frame, and then performs one or more of the following functions: (1) the source node removes its own transmitted frames that complete one trip around the ring through re-circulation; (2) once the
destination address of the LAN traffic matches the node’s MAC address, it is copied and delivered to the end users; (3) all US traffic (including LAN and control frames), excluding those that match items 1 and 2 above, is processed, regenerated, and then retransmitted to the next node.

Figure 3-1: Standalone 10G-EPON Architecture
3.2.2 Protected State Architecture

The protected architecture as shown in figure 3-3 is identical to that of the normal working architecture except for the following additional components: i) a redundant trunk fiber and distribution fiber ring; ii) a redundant transceiver pair located at the OLT; iii), Automatic Protection Switching (APS) module located at each ONU. The APS module attached to each ONU monitors the state of its adjacent distribution fiber paths and the state of the ONU and performs both fault detection and the APS functions. Each APS module houses a commercially available low loss 4x4 bidirectional Optical Switch (OS) that is capable of switching from any port to any port used for switching between working and protection fibers. It also includes two detection circuits comprised of a 1×2 CWDM filter (to separate the combined DS/US signal), a control circuit to configure the OS, and a p-i-n detector (except the first ONU (ONU1), which has two p-i-n detectors at the first detection circuit). The first detection circuit of each ONU (except the first ONU) is used to detect only the US signal via taping a small portion (about 1%) of the incoming combined (DS/US) signal and passing it through the CWDM filter. On the other hand,
the first detection circuit of the first ONU is used to detect both US and DS signals. Likewise, the second detection circuit of each ONU is used to detect the outgoing US signal via taping a small portion (about 1%) of the outgoing combined signal.
Figure 3-3: 10G EPON Protection State Architecture
3.3 Recovery Time

Recovery time is defined here as the time elapsed from when a failure occurs to when service is fully restored and a new cycle resumes. The total recovery time is the sum of several delay components including timeout, fault detection time, REPORT/GATE transmission time/propagation delays/processing times, and OS switching time. In general, the switching time is much longer than all other delay components combined and, therefore, the total recovery time is mainly dominated by the switching time (about 13 ms) [1].

3.4 Power Budget and scalability Analysis

The scalability of the proposed working state architecture is mainly limited by the concatenated splitter losses encountered by the DS signal at each node. Since the US signal is regenerated at every node, typical limited US power budget problems as well as the utilization of the 10 Gbps US burst-mode Tx/Rx and associated design challenges at the ONU/OLT are totally eliminated. To examine the performance impact of the DS power budget under the assumption of a fixed (10:90) tap ratio at each ONU, we consider the worst-case scenario by calculating the total ODN loss (passive optical elements (e.g., splitters, combiners, fibers, connectors, switches and splices forming an optical path), incurred by the DS signal on its optical path from the OLT to the second to last ONU (ONU_{N-1}).

There are two types of losses encountered by the DS signal at each node. The first type is along the path I-A-B in Fig. 3-3 (Drop-component, $IL_{Drop}$) and the second type is along the path I-A-E-O (Express-component, $IL_{Express}$). Table I quantifies both types of losses assuming typical commercially available CWDM components.
The total ODN loss incurred by the downstream signal on its path to ONU_{N-1} is:

\[
IL_{\text{ONU}_{N-1}}^{\text{Total Loss}} = IL_{\text{fiber}}^{\text{trunk}} + 2 \cdot IL_{\text{CWDM}} + (N - 2) \cdot IL_{\text{Express}}^{\text{ONU}} + IL_{\text{Drop}}^{\text{ONU}} + IL_{\text{Ring}}^{\text{fiber}},
\] (1)

Assuming a 20 km trunk feeder fiber (0.25 dB/km loss), the first ONU is 20 km away from the OLT, and the last ONU is 23.2 km away from the OLT (ring circumference is about 3.2 km; 1 km diameter), and the IEEE 802.3av 10G-EPON highest power budget class (PR/PRX30) parameters [2] with a DS Rx (APD w/FEC) sensitivity of −28.5 dBm and OLT Tx optical power of +2 dBm, the total number of ONUs that can be adequately supported is equal to 10 ONUs, (see Fig. 3-4). As for the protected state architecture, the signals encounter the additional OS and tap loss at each node. Assuming a 0.5 dB insertion loss per OS, the total number of ONUs that can be adequately supported by the protected architecture is reduced to 7 ONUs shown in Fig.3-5.
**10 Nodes Network Architecture**

1. **Express Path**
   -1.6 dB per node for express path to next node
2. **Drop Path**
   -10.5 dB loss
3. **Fiber Ring**
   -0.8 dB loss

---

**Figure 3-4: 10 Nodes Network Architecture**
3.5 Experiment

The transmission performance of the ring-based passive optical network is measured at 10Gbps data rate (downstream) and compared to the results collected at 2.5 Gbps. The major performance metrics used are typical BER and eye diagram.

The experimental setup is shown in Figure 3-7. Since there is no 10:90 output coupler in last ONU (see Figure 3-6), the ONU before last will receive the lowest, and usually the worst, signals. The received power of the ONU before last is about 9~10dB smaller than that of last ONU, therefore the measurement of BER and eye diagram is conducted at the ONU before last.
Figure 3-6: Inside structures of (a) Distribution node (b) ONU{s excluding the last one and (c) last ONU

Figure 3-7: Experimental setup for the measurement of BER of the ONU before last under 10Gbps downstream communication

10Gbps PRBS signals from BERT (Agilent N4901B 13.5Gb/s serial BERT) were sent to RF driver to drive a MZ amplitude modulator (SDL 10Gb/s amplitude modulator). Then the modulated optical signals were amplified by EDFA, passed through a variable optical attenuator
(VOA) and finally reached the ring. The output power after the EDFA is 7dBm. There are 10 ONUs in the ring.

The measured bit error rate and eye diagram at -17.5dBm received power are shown in Fig.3-8.

![Figure 3-8: Measured (a) bit error rate under different received power (b) eye diagram at -17.5dBm received power of the ONU before last](image)

The experimental results indicate that the system can adequately support up to 10 ONUs in the downstream direction under 10Gbps data rate.

Compared with the BER of same structure at 2.5Gbps as in Figure 3-9, the received power has to increase 9~11dB to provide same BER. The slope of the curves in Figure 3-9 are more steep than that in Figure 3-8, which might be attributed to the difference of signal source in two experiments.
Figure 3-9: Measured bit error rate of same structure at 2.5Gbps
Chapter 4

4 EPON-Based 4G Mobile Backhaul RAN

Architecture
4.1 Introduction
To illustrate some of the key technical challenges associated with devising an all-packet-based 4G mobile backhaul RAN architecture that is built on top of a typically centralized PON infrastructure, it is first important to understand the novel and radical changes associated with the evolving 4G LTE/LTE-A RAN architecture [1-3]. First, the 3G-RNC is eliminated from the data path and its typical functions are incorporated into the eNB, including all radio control functions such as radio resource management, handover control, admission control, etc. Thus, the distributed nature of the LTE RAN architecture calls for new radio control algorithms and procedures that operate in a distributed manner. Second, the X2 protocol within LTE is a means for LTE BSs to exchange signaling and control messages (for user handover) directly with each other rather than having to send traffic up to a more centralized controller (e. g., EPC). Thus, backhaul architecture must be designed to facilitate rapid and reliable inter-BS local communications via selecting a short route between adjacent BSs (to lower the X2 delay) rather than the going back to the mobile core (much longer route). The LTE-A CoMP techniques imposes further constraint on the X2 delay (should be in a range of 1 ms or lower [4]).
Third, with RNC functionality distributed to the eNBs, LTE creates a requirement for fully meshing the eNBs – some 10,000 to 40,000 for a mobile operator running a network in a ‘mature’ market. The cause is call handoff. A call originating on the ‘anchor’ eNB that subsequently travels across the country - moving from eNB-to-eNB requires that each eNB utilized will need to communicate with both the anchor eNB (e.g., for billing purposes) and the previous and next eNB for call handoff [7]. Even though IP-based virtual private networks (VPNs) can be considered good candidates for this type of meshing, however, to support a mesh of this size, some degree of hierarchy is likely to be required. A hierarchical IPVPN may provide an ideal migration path for LTE and the all-IP RAN. To realize the full potential of LTE,
mobile operators must deploy such a kind of networking architecture in the future; however, the complexity as well as the highly prohibitive cost of such a large-scale deployment might make its implementation rather a distant future.

4.2 EPON-Based 4G Mobile Backhaul RAN Architecture and/or Converged Fixed Mobile Optical Access Network Architecture

The standalone ring-based EPON architecture can be evolved to an all-packet-based converged fixed-mobile optical access networking transport infrastructure (or just 4G mobile backhaul RAN), as shown in Figure 4.1, by simply interconnecting (overlying) the ONUs with the 4G’s BSs (WiMAX or LTE) and the OLT with LTE’s access gateway (AGW) or WiMAX’s access service node (ASN). Under this simple overlay (independent) model, the PON and 4G systems are operated independently where the RAN system is assumed to have its own NCM operations, independent of those for the PON. The BS is assumed to be collocated with an ONU or treated as a generic user attached to it. The ONU and BS can be interconnected as long as they support a common standard interface. Thus, the OLT, ASN/AGW, ONUs, and 4G BSs, are all assumed to support a common standard interface (e.g., 802.3ah Ethernet interface). Each ONU is assumed to have two different Ethernet port ranges; the first port range will support wired users, while the second port range will support mobile users. The port ranges will be used by the ONUs to identify and differentiate between mobile users versus fixed users.

There are two interconnection models (depending on how the ONUs are interconnected to the BSs), namely, the overlay (independent) model and the integrated model [5-6]. Under the simple overlay (independent) model, the PON and LTE/WiMAX systems are operated independently by considering an LTE/WiMAX BS a generic user attached to an ONU and/or collocated with it. The RAN architecture is assumed to have its own NCM operations, independent of those for the
PON architecture. Under the integrated model, an ONU and a BS can be integrated into a single module in terms of either software or both software and hardware functionalities. Since EPC aggregate traffic from thousands of eNBs, numerous OLTs can be attached to it (only two are shown in Figure 4-1 for simplicity).

Under the integrated model, an ONU and LTE’s BS (eNB) or WiMAX BS can be functionally integrated into a single module either in terms of software or both software and hardware functionalities. The following are the main technical requirements needed to support the functional integration of the PON and 4G LTE/WiMAX access infrastructure: 1) the OLT, S-GW/ASN, ONUs, and eNBs/BSs, are all assumed to support a common standard interface (e.g., 802.3ah Ethernet interface); 2) each ONU is assumed to have two different Ethernet port ranges, the first port range will support wired users, while the second port range will support mobile users. The port ranges will be used by the ONUs to identify and differentiate between mobile versus fixed users; 3) Depending on the selection of either Layer-3 or Layer-2 connectivity at the transport layer, all the intermediate nodes (e.g., OLT, ONU, eNB) in Figure 4-1 are assumed to be equipped with either an IP access router to forward IP packets or GE Ethernet switch to forward the traffic using either PBB, PBT, or VPLS. This is a critical issue that will be thoroughly investigated during the course of this work.

Figure 4-2b illustrates the three main control modules of the functionally integrated ONU-eNB access node, namely, ONU’s control module, eNB’s control module, and the common control module, where each module can be a single CPU in hardware [5]. ONU’s module interfaces with the PON section and runs the PON protocols; eNB’s module interfaces with LTE section and runs the LTE protocols. The common module interfaces to both the PON and eNB sections, manages and coordinates joint optical-radio resources, and executes the integrated DBA and
packet scheduling algorithms. ONU and eNB modules report their queue statuses and bandwidth request details to the common module; the latter utilizes this information to make decisions, and to optimally allocate upstream/LAN resources to the ONUs and eNBs.

The functional modules for provisioning upstream traffic corresponding to the three modules in Figure 4-2b are shown in Figure 4-2a. Specifically, the ONU’s control module that interfaces with the PON section includes the functional components of PON packet scheduler, priority queues management, and PON packet classifier. Similarly, the LTE’s module that interfaces to the LTE section includes the functional components of two LTE mapping modules (one to map UE’s radio bearers to mobility tunnels), eNB packet classifier, and LTE upstream scheduler. Finally, the third module at the bottom of

Figure 4-2: (a) Architecture of the ONU-eNB, (b) Functional Modules hardware layout
Figure 4-2a corresponds to the ONU-eNB common coordinator controller, which comprises the functional components required to map QoS between PON and LTE and performs global admission and congestion control as well as integrated DBA and resource allocation and sharing protocols and algorithms.

Similarly Figure 4-3b illustrates the three main control modules of the functionally integrated ONU-BS access node, namely, ONU’s control module, WiMAX BS’s control module, and the common control module, where each module can be a single CPU in hardware. The functional modules for provisioning upstream traffic corresponding to the three modules in Figure 4-3b are shown in Figure 4-3a.

![Figure 4-3 (a)](image1)

![Figure 4-3 (b)](image2)

Figure 4-3: (a) Architecture of the ONU-BS, (b) Functional Modules hardware layout
Chapter 5

5 Key Enabling networking Innovations
This section introduces the key networking building blocks that enable the standalone ring-based EPON architecture to support a fully distributed 4G mobile backhaul RAN architecture and/or converged/unified fixed-mobile optical access networking infrastructure. Specifically, we will develop efficient methodologies to address the key critical enabling networking building blocks including: 1) A Fully Distributed Control Plane; 2) Dynamic Bandwidth Allocation; 3) QoS Support & Mapping; and 4) Layer-2 versus Layer-3 connectivity at the transport layer.

5.1 Fully Distributed Control Plane

This work utilizes the control and management messages defined by the IEEE 802.3ah multi-point control protocol (MPCP) standard [10] that facilitate the exchange of control and management information between the ONUs/eNBs and OLT and between the ONU|BSs. The protocol relies on two Ethernet control messages, GATE (from OLT to ONUs/eNBs) and REPORT (from ONUs|BSs to OLT and between ONUs/BSs) messages in its regular operation. Direct communication among ONUs/eNBs is achieved via the US wavelength channel (control messages along with both LAN and US data share the same US channel bandwidth (in-band signaling)), which is terminated, processed, regenerated, and retransmitted at each ONU.

Since control messages are processed and retransmitted at each node, the ONUs can directly communicate their US/LAN queue status and exchange signaling and control information with one another in a fully distributed fashion. Likewise, BSs can also directly communicate the status of their queues and radio resources and exchange signaling and control messages with one another. The control plane utilized among the ONUs/BSs can thus support a distributed PON-4G RAN architecture, where each access node (ONU/BS) deployed around the ring has now a truly physical connectivity and is, thus, capable of directly communicating with all other access nodes, in conformity with 4G LTE standards.
Each access node maintains a database about the states of its queue and the state of every other ONU/BS’s queue on the ring. This information is updated each cycle whenever the ONU receives new REPORT messages from all other ONUs. During each cycle, the access nodes sequentially transmit their REPORT messages along with both US and LAN data in an ascending order within their granted timeslots around the ring from one node to the next, where each REPORT message is finally removed by the source ONU after making one trip around the ring. The REPORT message typically contains the desired size of the next timeslot based on the current ONU’s buffer occupancy. Note that the REPORT message contains the aggregate bandwidth of both fixed and mobile data buffered at each ONU’s/eNBs’s queue (requested size of next timeslot).

Since the US channel is shared among all ONUs/eNBs, a distributed DBA scheme is required to efficiently and fairly provision US/LAN traffic among the ONUs/eNBs. The control plane utilizes a time division multiple access (TDMA) arbitration scheme to implement fully distributed DBA and packet scheduling algorithms in which the OLT/AGW is excluded from the arbitration process. It assumes a cycle-based upstream link, where the cycle size can have fixed or variable length confined within certain lower and upper bounds to accommodate the dynamic upstream traffic conditions. Note that under normal operation, all ONUs are synchronized to a common reference clock extracted from the OLT’s downstream traffic. The synchronization scheme is necessary for the execution of the distributed DBA scheme.

An identical dynamic bandwidth allocation (DBA) module, which resides at each access node (ONU/ eNB), uses the REPORT messages during each cycle to calculate a new US timeslot assignment for each ONU. ONUs sequentially and independently run instances of the same DBA algorithm outputting identical bandwidth allocation results each cycle. The execution of the
algorithm at each ONU starts immediately following the collection of all REPORT messages. Thus, all ONUs must execute the DBA algorithm prior to the expiration of the current cycle so that bandwidth allocations scheduled for the next cycle are guaranteed to be ready by the end of the current cycle. An execution of the DBA algorithm produces a unique and identical set of ONU/BS assignments. It is critical that the algorithm produces a unique outcome for any arbitrary set of inputs. Once the algorithm is executed, the ONUs sequentially and orderly transmit their data without any collisions, eliminating the OLT's centralized task of processing requests and generating grants for bandwidth allocations.

Thus, supported by the distributed control plane, most of the typical radio control functions including radio resource management, handover control, admission control, etc, can be independently implemented at each eNB in a distributed approach without resorting to a central control entity (e.g., AGW), in conformity with 4G LTE standards. Likewise, most of the typical wireline control functionalities including DBA, queue management, packet scheduling, and restoration algorithms, can be independently implemented at each ONU in a distributed approach without resorting to a central control entity (e.g. OLT). These functionalities are typically implemented at the distant OLT/RNC in today’s standalone centralized PON/RAN systems.

Please note that under the overlay model, the distributed control plane enables each BS attached to the ring to independently provision only mobile traffic (upstream and LAN traffic). Likewise, it also enables each ONU to independently provision only fixed traffic but not both (since the third common control module is absent). However, under the integrated model, the inclusion of the third common control module enables the support of a unified control and management plane
that manages and controls the overall fixed optical and mobile radio network resources. Thus, the control plane now is both distributed and unified. Supported by such a unified distributed control plane, the processes of bandwidth allocation, packet scheduling, and queue management for both the PON and RAN systems are integrated and operated in a distributed manner.

Thus, the integrated model provides the best overall system performance in terms of cost-effectiveness, bandwidth utilization, and support for better QoS. This is because the integrated control module has global information about the entire fixed/mobile network status including the aggregate bandwidth requirements of both wired and mobile users. Hence, the processes of bandwidth allocation and packet scheduling as well as prioritizing different class of services (for either wired or mobile users) are optimized throughout the network system.

5.2 Fully Distributed Dynamic Bandwidth Allocation at the ONU/BSs

5.2.1 Overview of Typical PON Scheduling Schemes

In order for mainstream centralized EPON architectures to support differentiated QoS, two independent scheduling mechanisms are required:

a) Scheduling at the OLT (inter-ONU scheduling): The OLT is the only device that can arbitrate the upstream transmissions by allocating an appropriate TW to each ONU. In this case, the OLT passes the request messages to a dynamic bandwidth allocation module (co-located with the OLT) that performs the bandwidth allocation computation and generates grant messages.

b) Scheduling at the ONU (intra-ONU scheduling): In this case, queue management and priority queuing are used to divide the bandwidth allocated by the OLT to a given ONU among the different class of services (based on their priorities) supported by that ONU.
Since the two scheduling schemes are independent of each other, the final bandwidth allocated to a particular class of service for a given ONU may not be the optimum choice.

Several centralized tree-based DBA schemes have recently been reported in the literature [9-12]. An OLT-based polling scheme, called \textit{Interleaved Polling with Adaptive Cycle Time} (IPACT) based on \textit{Grant} and \textit{Request} messages, has been presented in [9]. Using IPACT, several DBA schemes were studied in [9]; namely fixed, limited, gated, constant credit, and linear credit. Amongst these algorithms, the limited was shown to exhibit the best performance. The limited DBA scheme is cycle-based, where a cycle ($T_{CYC}$) is defined as the time that elapses between two executions of the scheduling algorithm. A cycle has a variable length size confined within certain lower and upper bounds, which we denote as $T_{MIN}$ and $T_{MAX}$ (sec) respectively. Thus, the algorithm schedules between $B_{MIN}$ and $B_{MAX}$ (bytes) at a time, where $B_i$ is determined by multiplying $T_i$ with the line rate. In this scheme, the ONU will be granted the requested number of bytes, but no more than a given predetermined maximum $B_{MAX}$. If $R_i$ is the requested bandwidth of ONU$_i$, then the granted bandwidth ($B_{Granted}$) is equal to:

$$B_{Granted} = \begin{cases} R_i & \text{if } R_i \leq B_{MAX} \\ B_{MAX} & \text{if } R_i > B_{MAX} \end{cases}$$

(1)

$B_{MAX}$ is determined by the maximum cycle time $T_{MAX}$ [9]:

$$B_{max} = \frac{I}{N} [R_{EPON}T_{MAX} - (N*T_G)]$$

(2)

where $N$ is the number of ONUs, $T_G$ is the guard band slot, and $R_{EPON}$ is EPON line rate.
5.2.2 Decentralized Dynamic Bandwidth Allocation Scheme

All of the above referenced DBA schemes are OLT-based, that is the OLT has centralized intelligence. The performance of most of these centralized schemes, including the limited scheme, suffers from several limitations, including: (1) the bandwidth granted by the OLT, during cycle n, to ONU\textsubscript{i} is only determined by the content of a single REPORT message transmitted in the previous cycle n-1 by ONU\textsubscript{i} (i.e., the bandwidth computation module does not take into account the remaining requests of other ONUs). Thus, the process of bandwidth allocation is not globally optimized; (2) due to the bursty nature of Ethernet traffic, some ONUs might have less traffic to transmit while other ONUs may require more bandwidth than B\textsubscript{max}. For instance, assume that ONU\textsubscript{i} requests an amount of bandwidth R\textsubscript{i} < B\textsubscript{max}, while ONU\textsubscript{j} requests an amount of bandwidth R\textsubscript{j} > B\textsubscript{max}. Although there is an excess amount of bandwidth (B\textsubscript{max} - R\textsubscript{i}) that can be granted to ONU\textsubscript{j}, however, due to limitation # 1 cited above, the maximum bandwidth that may be granted to ONU\textsubscript{j} is only B\textsubscript{max}.

The proposed distributed ring-based EPON architecture, however, enables instances of the same DBA algorithm (inter-ONU scheduling) to be executed simultaneously at each ONU. Thus, both scheduling tasks (inter and intra-ONU scheduling) schemes are performed at the ONU leading to the notion of integrating both scheduling mechanisms at the ONU.

In this work, the centralized limited service scheme reported in [9], along with the appropriate changes needed to accommodate the distributed architecture, is used here as the basis for the decentralized DBA scheme presented here. As mentioned above, to globally optimize the bandwidth allocation process, the proposed DBA algorithm execution is performed only after each ONU receives and processes all other ONUs requests.
Based on bandwidth demands, ONUs can be classified into two groups, namely: lightly loaded ONUs that have bandwidth demands less than $B_{MAX}$; and heavily loaded ONUs that have bandwidth demands more than $B_{MAX}$. During each cycle, the DBA module must keep track of the unclaimed bandwidth from the set of lightly loaded ONUs. It then must redistribute this excess bandwidth to other heavily loaded ONUs based on their requested bandwidth, i.e. two ONUs requesting bandwidths $B_1$ and $B_2$ more than $B_{MAX}$ will be assigned excess bandwidths proportional to $B_1$ and $B_2$.

During each cycle, the lightly loaded ONUs with $R_i < B_{MAX}$ will contribute a total remainder cycle bandwidth:

$$B_{Cycle\_Remainder} = \sum_{i=1}^{L} (B_{MAX} - R_i) \quad L: \text{Number of lightly loaded ONUs}$$

The heavily loaded ONUs with $R_i > B_{MAX}$ will require a total over the limit cycle bandwidth:

$$B_{Cycle\_OverLimit} = \sum_{i=1}^{H} (R_i - B_{MAX}) \quad H: \text{Number of heavily loaded ONUs}$$

The total remainder cycle bandwidth can be fairly distributed amongst the heavily loaded ONUs to expand their maximum transmission window as follows [10]:

$$\Delta B_i^{extra} = B_{Cycle\_Remainder} \left[ \frac{R_i - B_{MAX}}{B_{Cycle\_OverLimit}} \right]$$

(3)

where $\Delta B_i$ is the extra bandwidth allocated to ONU$_i$. The granted bandwidth, $B_{GH}$, for a heavily loaded ONU$_i$ is given by:

$$B_{GH} = \Delta B_i^{extra} + B_{MAX}$$

(4)

If $R_i$ is the requested bandwidth of ONU$_i$, $B_{Granted}$ is the bandwidth granted using the proposed limited service-based distributed DBA scheme (Eqs. 1 and 4), then $B_{Granted}$ can be expressed as:
\[
B_{\text{Granted}} = \begin{cases} 
R_i & \text{If } R_i \leq B_{\text{MAX}} \\
R_i & \text{If } R_i > B_{\text{MAX}} \land B_{\text{Cycle Remainder}} \geq B_{\text{Cycle OverLimit}} \\
B_{\text{GH}} & \text{If } R_i > B_{\text{MAX}} \land B_{\text{Cycle Remainder}} < B_{\text{Cycle OverLimit}} 
\end{cases}
\]

Note that the lightly loaded ONUs (\(R_i < B_{\text{MAX}}\)) can be scheduled instantaneously “on-the-fly” without waiting for DBA module to perform its end of cycle computations. Whereas, the heavily loaded ONUs (\(R_i > B_{\text{MAX}}\)) will have to wait until all REPORT messages have been received and the DBA algorithm has computed their bandwidth allocations. Thus, lightly loaded ONUs can be scheduled ahead of heavily loaded ones.

Thus, the proposed decentralized EPON architecture addresses some of the limitations of the centralized DBA schemes cited above and can further provide several advantages as follows:

- Since the bandwidth allocation computation is performed after receiving and processing all ONUs requests (processing period) (i.e., the computation takes into account the entire network status), the bandwidth allocation process now reflects the entire network information collectively, leading to a globally optimized decision.

- In contrast to the centralized architectures where the order of ONUs transmission is fixed in each cycle (sequential), the decentralized architecture has the added flexibility of varying the order of ONUs transmission according to the ONUs traffic demands and priority. Thus, the order of ONUs transmission may be different in each cycle and need not be fixed.

- Since the DBA computation is based on the global network information, the heavily loaded ONUs may be allocated the remaining excessive bandwidth that is not utilized by the lightly loaded ONUs.
Given that DBA and priority queuing scheduling tasks are both executed at the ONU, the DBA module can integrate both scheduling information to yield a globally optimized bandwidth allocation to a particular class of service in a given ONU.

Because the centralized limited DBA scheme was shown to exhibit the best performance in [7-8], we will consider this scheme as a reference model for comparing the performance of our distributed architecture versus that of the centralized scheme reported therein.

5.3 QoS Mapping

A. Mobile WiMAX

Typical mobile WiMAX MAC is centralized and connection-oriented. A connection identifier (CID) identifies each WiMAX connection. The main mechanism for providing a connection-based QoS is to classify and associate packets traversing the MAC interface to IP Service Flows (SFs), where each existing SF is identified by a 32-bit SF identifier (SFID) and is characterized by a set of QoS parameters. A CID is then mapped to an SFID provided that the SF has already been admitted (active SF). Once the MS’s CIDs are terminated at the BS, they are mapped into the appropriate mobility tunnels based on their CIDs. The BS’s packet classifier then maps their constituent IP SFs into their appropriate priority queues based on CIDs attached to the IP packets. To allow for traffic separation in the PON-based transport network (IP cloud connecting the BSs to the OLT/ASN), the BS maps each CID onto a corresponding DiffServ Code Point (DSCP) in order to translate CID to transport-based QoS (DSCP). Using this mapping function,
packets on a given CID associated with specific QoS parameters are marked with a specific DSCP for forwarding in the transport network. The MPC performs the mapping for DL packets.

On the other hand, EPON technology does not support this type of CID-based connection. Rather, it supports only enhanced QoS through prioritization where packets are classified, stored in different priority queues and, then, scheduled for service according to their priority. In a typical centralized EPON, QoS support is implemented via two independent scheduling mechanisms [10]: 1) inter-ONU scheduling: an aggregate bandwidth is allocated to each ONU by the OLT. 2) intra-ONU scheduling: each ONU makes a local decision to allocate the granted bandwidth and schedules packets transmission for up to eight different priority queues in the ONU. In the case of the proposed architecture, however, instances of the same DBA algorithm are executed simultaneously at each ONU. Thus, both scheduling mechanisms (inter and intra-ONU scheduling) are performed at each ONU-BS in a fully distributed approach, leading to the notion of integrating both scheduling mechanisms at the ONU. This enables the proposed distributed architecture to provide better QoS support and guarantees.

For simplicity, we assume that each ONU maintains three separate priority queues that share the same buffering space. We consider three priority classes P0, P1, and P2, with P0 being the highest priority and P2 being the lowest. These classes are used for delivering voice (CBR), video stream (variable-bit-rate or VBR), and best-effort (BE) data, respectively, as they allow easy mapping of DiffServ’s Expedited Forwarding (EF), Assured Forwarding (AF), and BE classes into 802.1D classes. Since both EPON and mobile WiMAX classify data traffic in a differentiated services mode, an effective mapping mechanism is required between EPON priority queues and CID-based WiMAX IP flows. Specifically, the mapping has to identify which WiMAX IP flow should be stored in which EPON priority queue for equivalent QoS.
EPON has up to eight different priority queues in each ONU, while mobile WiMAX supports five classes of service including Unsolicited Grant Service (UGS), extended real-time Polling Service (ertPS), real-time Polling Service (rtPS), non real-time Polling Service (nrtPS), and Best Effort (BE). In this work we assume that WiMAX UGS queues are mapped into EPON P0 queue, rtPs into P1, and BE into P2.

B. LTE

The QoS model of LTE, which was standardized in 3GPP release 8, is based on the logical concept of an “EPS bearer” [1-3]. The term “bearer” refers to a logical IP transmission path between the UE and the EPC with specific QoS parameters (capacity, delay, packet loss error rate, etc.). Each bearer is assigned one and only one QoS class identifier (QCI) by the network and is composed of a radio bearer and a mobility tunnel. The QCI is a scalar that is used within the access network to identify the QoS characteristics that the EPC is expected to provide for the IP SDFs. This scalar (bearer ID) is used by routers to access node-specific parameters that control packet forwarding treatment (e.g., scheduling policy, admission thresholds, link layer configurations, queue management policy, etc.), which are specified and preconfigured by the operator. An EPS bearer uniquely identifies packet flows that receive the same packet forwarding treatment between the UE and EPC.

The 3GPP specifications define eight standardized QCIs, each with its corresponding standardized characteristics including bearer type (GBR versus non-GBR), priority, packet delay, and packet-error –loss rate. To allow for traffic separation in the transport network (IP cloud connecting the eNBs to the EPC), P-GW and eNB map each QCI onto a corresponding diffserv code point (DSCP) in order to translate a bearer-based QoS (QCI) to transport-based QoS
(DSCP) [1-2]. Using this mapping function, packets on a bearer associated with a specific QCI are marked with a specific DSCP for forwarding in the transport network. The QCI to DSCP mapping is performed based on operator policies, which are configured into the network nodes. P-GW performs the mapping for DL packets while eNB performs it for UL packets.

As can be seen from the eNB module shown on Figure 5-1, the UE uses the packet filters to classify IP packets to authorized IP SDFs. This process is referred to as SDF detection. The UE then performs the binding of the detected uplink IP SDFs to the appropriate bearers. Once the UE’s radio bearers are terminated at the eNB, they are mapped into the appropriate mobility tunnels based on their bearer-IDs. The eNB’s packet classifier then maps their constituent IP flows into their appropriate priority queues based on the bearer-IDs attached to the IP packets, which is the basic enabler for traffic separation. Finally, to allow for traffic separation in the transport network, the eNB maps each OCI (bearer-ID) onto the corresponding DSCP value.

On the other hand, EPON technology does not support this type of bearer-based connection. Rather, bandwidth requests are queue-oriented; an aggregate bandwidth is allocated to each
ONU, and then the latter makes a local decision to allocate the granted bandwidth and schedules packets transmission for up to eight different priority queues in the ONU. In a typical centralized EPON, QoS support is implemented via two independent scheduling mechanisms [10]: 1) inter-ONU scheduling: an aggregate bandwidth is allocated to each ONU by the OLT. 2) intra-ONU scheduling: each ONU makes a local decision to allocate the granted bandwidth and schedules packets transmission for up to eight different priority queues in the ONU. Under the proposed integrated architecture, however, instances of the same DBA algorithm are executed simultaneously at each ONU. Thus, both scheduling mechanisms (inter and intra-ONU scheduling) are performed at each ONU in a fully distributed approach, leading to the notion of integrating both scheduling mechanisms at the ONU. This enables the proposed distributed architecture to provide better QoS support and guarantees.

Both EPON and LTE classify data traffic in a differentiated services mode. However, EPON supports only enhanced QoS through prioritization where packets are classified, stored in different priority queues and, then, scheduled for service according to their priority. On the other hand, LTE supports guaranteed QoS through logical bearer reservation where each router/node on the RAN/EPC is configured to forward the packets of different IP flows based on their bearer-IDs (QCIs) in which resources are reserved (queue space, queuing management strategy, scheduling strategy) accordingly.

To achieve a truly integrated model, an effective mapping mechanism is required between EPON priority queues and QCI/bearer-based LTE IP flows. Specifically, the mapping has to identify which LTE IP flow should be stored in which EPON priority queue for equivalent QoS. EPON has up to eight different priority queues in each ONU, while LTE defines eight standardized QCIs that classify data traffic into eight different classes of service, ranging from real-time
gaming to the lowest priority best effort TCP bulk data. This theoretically facilitates a one-to-one mapping from eNB’s eight priority queues to ONU’s eight priority queues (e.g., packets of highest/lowest eNB’s priority queue are mapped onto highest/lowest ONU’s priority queue) and vice versa in both upstream and downstream directions. However, devising an efficient viable mapping strategy that enables a unified QoS model for both wired and wireless services requires the implementation of the following critical functions:

1) Since the bearers are not visible to the ONUs/OLTs, each and every ONU/OLT must be directly configured (semi-statically) with all eight LTE’s standardized QCIs (QoS characteristics) or more precisely with the corresponding DSCP values (QCI to DSCP mapping is performed based on operator policies). This configuration enables each ONU/OLT to forward the packets of different UL/DL IP flows based on their DSCP values such that the packets-forwarding treatment received by these flows at the ONU/OLT is identical to that received at the eNB/P-GW. This is achieved by ensuring that the queue management schemes and scheduling algorithms implemented at the ONU/OLT are identical to those implemented at the eNB/P-GW.

2) The PON’s packet scheduler at the ONU/OLT must apply the same packet forwarding treatment for both wired and wireless upstream/downstream traffic for each and every configured QCI/DSCP value that is associated with a given IP flow. This further enhances the typical PON’s prioritization-based QoS support for wired users as well as simplifies the implementations of queue management schemes and scheduling algorithms at the ONUs and OLTs.

3) The typical PON’s cycle-based approach for DBA and QoS support must be drastically modified at both the ONUs and OLTs. None of EPON scheduling mechanisms can guarantee bandwidth for real-time IP flows because the bandwidth allocated by the OLT to one ONU can only be guaranteed for a significantly short time (e.g., a fraction of one cycle) and may vary from
one cycle to another cycle according to the load at other ONUs. Thus, each ONU is required to reserve bandwidth for its real-time IP flows for the whole duration of the flow (and not on a per cycle basis) in order to satisfy their QoS requirements as specified by the attached DSCP value.

4) In addition to bandwidth allocation and service differentiation, a global admission and congestion control (AC) mechanism for both wired and wireless traffic that makes decisions on whether or not to admit/block a new wired/wireless real-time IP flow based on its requirements and the upstream channel usage condition. Ideally, this AC module should be housed at the common control module (Figure 5-1) since the critical information needed by the AC module to make appropriate admission/denial decisions (e.g., available fixed optical and mobile radio resources as well as both available wired (ONUS-OLT) and wireless (UE-eNBs) uplink channel capacities) is always dynamically available to the common control module. For instance, when the congestion bottleneck is at the backhaul and not at the radio interface, the common control module can block the admission of any new mobile user’s traffic until congestion subsides.

The combination of a distributed PON-RAN architecture along with a fully distributed/unified control plane with global information about the entire fixed-mobile network status collectively enable the implementation of a simple and efficient QoS-aware DBA scheme, in which resources are reserved (e.g., queue space and bandwidth) via signaling. Note that the overall process of QoS mapping and support can be further simplified by reducing the number of standardized QoS levels for both PON and LTE from eight to the typical three DiffServ’s classes of services (Expedited Forwarding (EF), Assured Forwarding (AF), and Best effort (BE)), which are commonly and widely used by operators.
5.4 Layer-2 versus Layer-3 Connectivity at the Transport Layer

Determining the most effective and efficient mix of layer-2 and layer-3 in the backhaul is a major issue worldwide. There are myriad approaches to support LTE backhaul. First, the transport network choice and architecture could have a significant impact on EPC – for instance, where layer-2 carrier Ethernet is utilized, a more centralized EPC gateway (S-GW and P-GW) is preferred. Alternatively, if layer-3 dominates, the gateway can be distributed and perhaps integrate that capability with edge routers. The EPC initial deployment is expected to be one of dedicated mobility nodes (S-GW and P-GW) installed on top of IP/MPLS core networks. However, over the longer term, there is a potential for eliminating the boundaries between the IP network and EPC “mobility layer” [11]. Under this scenario, we can assume that EPC applications can be implemented on a router where a dedicated module or blade is added to the router to provide EPC functions, resulting in a “Carrier-grade edge-router” that performs both typical routers and EPC functionality.

Thus, under the assumption of layer-3 connectivity, the EPC is modeled as a distributed architecture by pushing the S-GW and P-GE nodes to the edge and assume that these nodes are multiservice carrier-grade edge-routers incorporated with the typical IP/MPLS core network. On the contrary, the second assumption is that rather than pushing layer-3 routing and S-GW and P-GW out towards the edge of the network, the focus instead should be on low-cost layer-2 Ethernet transport, backhauling traffic to a more centralized S-GW and P-GW that are implemented in a blade server-platform or other non-router platforms [11]. Under this assumption (layer-2 Ethernet connectivity), several hierarchical carrier grade Ethernet transport solutions can be utilized including: a) the IEEE 802.1ad (Qin Q or “double tagging”); b) the
IEEE 802.1ah (MAC-in-MAC or Provider Backbone Bridges (PBB); c) Provider Backbone Transport (PBT), or PBT-TE (PBT with traffic engineering).

To avoid the complexity of IP/MPLS control planes, along with the fact that Ethernet is considered as the most effective method to transport IP packets, the LTE backhaul can, for example, use MEF-compliant interfaces on the eNB and on the S-GW and MME. The mobile operator can send VLAN-tagged frames toward the EPC. The backhaul can now identify the VLAN tag, then maps these frames to the EVCs (Ethernet virtual circuits). A multipoint EVC can be used to support X2 among a cluster of eNBs that need to exchange protocols. Initially, we lean more towards the approach of leaving the IP functionality to the mobile endpoints that actually need it (e.g., the eNB and the EPC), and avoiding it in the backhaul network by utilizing carrier Ethernet. A detailed technical and economic study is needed to weigh the pros and cons of each transport technology to determine the optimum solution.
Chapter 6

6 Key Technological and Economic Advantages Enabled By the Distributed EPON-Based RAN Architecture
The distributed ring-based architecture along with the supporting control plane enable the proposed EPON-based RAN architecture to support several salient networking features that collectively significantly augment the performance of both the RAN and MPC in terms of handoff capability, overall network throughput, stability, latency, and QoS support.

6.1 Enhanced Network Throughput and Stability

To illustrate the merits of the proposed distributed architecture, the performance of the proposed EPON-based 4G RAN (e.g., mobile WiMAX RAN) is compared with that of the typically centralized WiMAX RAN. Two simulation programs were developed using OPNET, one for the typical WiMAX RAN and the other one for the EPON-based RAN. The performance metrics used here are network utilization, handoff throughput, and end-to-end (ETE) delay, where ETE delay is defined here as the time that elapses between the times of arrival of packets at the MS queues and the time of their arrival at the OLT/ASN (US traffic) or at the TBSs around the ring (local mobile LAN traffic).

Two different traffic load scenarios are considered for both upstream and downstream traffic. In the first scenario, all BSs are assumed to have uniform traffic, i.e., all BSs have an equal average traffic load. Although this assumption is not realistic, this scenario is used only as a reference model. In the second scenario, where the significance of utilizing a packet-based RAN architecture is clearly established, we consider the most realistic and practical case of non-uniform traffic load in which, during a given period, some BSs might be lightly loaded/idle, while other BSs might be heavily loaded. At a given total network load, different BSs have different average traffic loads. This is achieved by dividing the BSs into “Super Heavily
Loaded”, “Heavily Loaded”, “Moderately Loaded” and “Lightly Loaded”. Table 6-1 shows the details for different load scenarios where every row corresponds to one traffic load (total network traffic load) scenario as well as to the different loads used in the following Figures (Fig. 1 through 6).

<table>
<thead>
<tr>
<th>Heavily Loaded BSs</th>
<th>Lightly Loaded BSs</th>
<th>Total Network Load</th>
</tr>
</thead>
<tbody>
<tr>
<td># BSs</td>
<td>BS Load</td>
<td># BSs</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>0.8488</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>0.8488</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>0.8488</td>
<td>11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Super Heavily Loaded BSs</th>
<th>Moderately Loaded BSs</th>
<th>Total Network Load</th>
</tr>
</thead>
<tbody>
<tr>
<td># BSs</td>
<td>BS Load</td>
<td># BSs</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>1.85</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>1.85</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>1.85</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>1.85</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 6-1: Unevenly Loaded Base Stations Scenarios

The following are the system parameters used for simulating the EPON-based RAN architecture:
(1) a PON with 16 ONUs, each serving varying number of BSs (a minimum of one BS to a maximum of 10 BSs), depending on the varying traffic load.; (2) aggregate access link data rate
from the MSs to a given ONU is 100 Mb/s; (3) the RAN DS line rate (from the OLT/ASN to the ONU/BSs) is assumed to be same as the US line rate (from the ONU/BSs to the OLT/ASN) and is equal to 1 Gb/s; (4) the average distance between the OLT/ASN and ONU/BSs is 20 km; (5) the buffer size in each ONU/BS is 1 Mbyte; (6) the maximum EPON cycle time is 2 ms for US transmission, while a standard fixed periodic cycle of 5 ms is assumed for WiMAX US transmission (from the MSs to the BS); (7) the IEEE 802.3ah MPCP REPORT/GATE message is 64 bytes; (8) we assume that all network traffic is just mobile traffic initiated by WiMAX’ MSs, i.e., traditional EPON’s fixed wired end-user’s traffic is assumed to be zero; (9) the total mobile traffic is divided equally among US mobile traffic and local mobile LAN; (10) the mobile traffic model used here is as described above in Section III B, with the three WiMAX CoSs (UGS, rtPs, and BE) are mapped into the three EPON CoSs (P0, PI, and P2), respectively; (11) the DBA scheme reported in [9] is used here to provision EPON US traffic, whereas the proportional fairness algorithm is used to provision WiMAX US traffic.

To have a fair comparison, all EPON-based RAN parameters listed above are also used for simulating the typical WiMAX except for the following: each and every dedicated link data rate of the typical WiMAX RAN in either US (16 dedicated point-to-point links between the ONU/BSs and the OLT/ASN) or DS (16 dedicated point-to-point links between the OLT/ASN and the ONU/BSs) direction is set to 62.5Mbps. Thus, the aggregated link data rate in either direction is: 62.5Mbps * 16 = 1 Gbps, which is equal to that of the EPON-based RAN.

Figures 6-1 and 6-2 show the uplink utilization versus time at a given single network load of 0.83 for evenly loaded and unevenly loaded BSs, respectively, for both the typical WiMAX and EPON-based RAN architectures. The results of both Figures demonstrate that EPON-based RAN has much higher utilization as well as stability with less variation with time compared to typical
WiMAX. This enhances the network’s stability and predictability. Figure 6-3 shows average uplink throughput versus total network load. Figure 6-4 shows the ETE network delay for both rtPs and BE traffic for both the typical WiMAX and EPON-based RAN architectures. It can be seen from the figure that at lower load, the typical WiMAX has lower delay that the EPON-based RAN. This is expected as typical WiMAX traffic utilizes dedicated point-to-point links and the queuing delay are still almost zero. However at higher traffic load, as expected, the EPON-based RAN exhibits much lower delay. Figure 6-5 and Figure 6-6 show uplink network utilization in case of evenly and unevenly loaded BSs respectively.

![Wireless Backhaul Time Series at Network Load of 0.83(Evenly Loaded BS's)](image)

Figure 6-1: Uplink Utilization Time Series for Evenly loaded BSs
Figure 6-2: Uplink Utilization Time Series for Unevenly loaded BSs

Figure 6-3: Uplink Ave Throughput
Figure 6-4: Average network ETE delay for rtPs and BE traffic

Figure 6-5: WiMAX Backhaul Uplink Utilization with evenly loaded BSs.
Figure 6-6: WiMAX Backhaul Uplink Utilization with unevenly loaded BSs.

6.2 Significance of Local Mobile LAN Traffic:

Local mobile LAN traffic is defined here as bidirectional multimedia traffic exchange (including VOIP, video, and data sessions) between two mobile users served by two different BSs that are either collocated or attached with/to two different ONUs on the same ring (same PON domain). In the proposed EPON-based RAN architecture, this traffic is directly routed on the ring from the source BS directly to the destination BS and vice-versa as local LAN traffic, without the direct participation of either the OLT or the MPC (e.g., ASN/AGW). This is significant as the volume of VOIP calls and/or multimedia data exchange between all local mobile users that are served by the many different BSs that are attached to the same ring is substantial. In a typical mobile WiMAX and/or LTE RAN, however, this traffic represents bidirectional US/DS data exchange between the two mobile users, which must be routed first from the source BS to the MPC (US traffic) and then from the MPC to the destination BS (DS traffic), and vice-versa.
Thus, a substantial volume of local mobile traffic and associated signaling overhead as well as the lengthy and complex processing of this traffic (e.g., sessions (LTE bearers/mobility tunnels) switch/set-up, retain, and tear-down and associated signaling commands from the BSs to the MPC and vice-versa) have been offloaded from the overburdened MPC to the access nodes (ONUS/BSs) of the RAN. This has a significant impact on the performance of the MPC. First, it frees up a sizable fraction of the badly needed network resources as well as processing on the centralized serving nodes (e.g. ASN/AGW) in the MPC to handle Internet-bound traffic more efficiently. Second, it frees up capacity and sessions on the typically congested mobile backhaul from the BSs to the MPC and vice-versa.

Figures 6-7 and 6-8 show the DS throughput with local upstream load of 0.5 and 1.0 respectively. In the typical WiMAX network, it is obvious that DS throughput is badly impacted as the local upstream traffic is increased. This is because the local upstream traffic is routed through the OLT/ASN and takes part of the DS traffic, hence impacting the DS throughput. On the other hand DS throughput of our proposed distributed EPON-based WiMAX network is independent of local traffic. Similarly Figures 6-9 shows actual US throughput with no local LAN traffic and Figure 6-10 shows US throughput when the local LAN traffic is increased to 0.5 in case of EPON-based and typical WiMAX architectures. Therefore as the local LAN traffic increases the US throughput is adversely impacted in case of typical WiMAX while in case of EPON-based WiMAX, US throughput remains independent of increasing local LAN traffic.

Finally, Figures 6-11 and 6-12 show average End-to-End (ETE) delay and packet drop respectively in case of EPON-based and typical WiMAX vs DS Load when one BS transmits high data rate (> 200 Mbps) to the other BS in the same ring( or in the same cell in typical WiMAX). Results show that packet ETE delay and packet drop rate increase as the DS load
increase in typical WiMAX while in case of EPON-based WiMAX ETE delay and packet drop remain independent of increasing DS load.

Figure 6-7 DS throughput with local load of 0.5

Figure 6-8: DS throughput with local load of 1.0
Figure 6-9: US Throughput with no local LAN Traffic

Figure 6-10: US Throughput with 0.5 local LAN Traffic
Figure 6-11: Average ETE Delay when one BS sends high data rate traffic to other BS in the EPON-based architecture as compared to typical WiMAX architecture vs DS Load.

Figure 6-12: Average ETE Delay when one BS sends traffic to other BS in the EPON-based architecture as compared to the typical WiMAX architecture vs DS Load.
6.3 Mobility Management - Enhanced Handoff Capabilities

In both mobile WiMAX and LTE standards hard handoff (HHO) is mandatory. The HHO is a break-before-make procedure, in which WiMAX mobile station (MS) and/or LTE user equipment (UE) breaks its connections with the serving BS (SBS) before setting up new connections with the target BS (TBS) and this is when traffic interruption and packet loss take place. By exploiting both the distributed nature of the ring-based RAN architecture and the supporting control plane, the proposed architecture enables the support of seamless and speedy inter-BS HOs in which, as the simulation results will show, packet loss is almost totally avoided and VoIP and other real-time IP applications can be adequately supported during HO. This is accomplished as follows:

1) When a MS/UE enters a domain served by the PON-RAN, it needs to register itself to the domain OLT’s access router and updates the new location in its home subscriber server (HSS). As long as the MS/UE is roaming within the same PON-RAN domain, it needs not to reregister again.

2) The physical connectivity among the both the SBS and TBS attached to the ring allows direct data exchange and intercommunications among them during HO (compare the simplicity and reduced latency and signaling overhead of this direct approach versus that of the typical 4G indirect bidirectional lengthy intercommunications and logical connectivity among the SBS and TBS via the MPC). Thus, once the TBS accepts the HO command, the SBS may immediately start to forward the buffered data (which have not yet been successfully sent to the MS), to the TBS directly on the ring as local LAN traffic. This is significant as creating the typical 4G logical connectivity among the SBS and TBS, which requires the lengthy process of signaling to
the ASN/AGW to coordinate the mobility-tunnel set up/switch from the SBS to TBS (and vice-versa) via the MPC, is totally avoided as well as the direct participation of the ASN/AGW/OLT.  

3) For the HO to complete, the TBS signals the OLT/ASN to inform it that the HO is complete and to update its records with the new TBS, i.e., to add TBS (and corresponding target ONU (TONU) that is collocated or attached with/to the TBS) to the forwarding list for the MS. Then, under the typical 4G RAN scenario, to resume normal operation and forward DS traffic to the TBS (or MS), the typical lengthy process of setting up a mobility tunnel form the MPC to the TBS is essential. Under the proposed PON-based RAN architecture, however, the scheduler at the OLT just simply redirects the MS’s DS traffic from the DS queue that was serving the SONU/SBS before the HO (the OLT houses N dedicated DS queues, each serving one of the N ONUs-BSs attached to the ring) to the new DS queue that is now serving the TONU/TBS. To further reduce the signaling latency and packet loss during the HO, the OLT may concurrently broadcast DS traffic destined to the MS to both the SBS and TBS.  

Overall, the proposed EPON-based RAN architecture introduces several significant advantages versus that of a typical mobile WiMAX and/or LTE RAN, including: 1) significantly reduces the signaling overhead and handoff latency; 2) offloading a sizable fraction of the local mobile sessions switch/set-up and tear-down and associated lengthy and complex signaling processing from the overloaded MPC to the RAN’s access nodes; 3) re-registration procedures to the HSS when the MS/UE moves from a BS to another is avoided as long as the MS/UE roams within the coverage area served by the BSs attached to the ring; 4) during inter-BSs HOs, no path switch/setup command is needed since the path (mobility tunnels) from MPC to the MS/UE remains unchanged.
6.3.1 IEEE 802.16e Handover Process Overview:

In general handover process consists of normal operation, cell reselection, handover decision, handover initiation and handover execution. The MS is connected to the SBS for the packet scheduling process. Periodic ranging takes place for the entire duration MS is connected to its SBS.

Cell Reselection or Scanning: It refers to the MS scanning and/or Association with one or more BSs. Base stations periodically broadcast network topology information using MOB_NBR-ADV message. The message includes the BSIDs of the neighboring BS’s along with their channel characteristics provided by each BS by the transmission of Downlink/Uplink Channel Descriptor (DCD/UCD) message. This enables MS to synchronize with the neighboring BSs by eliminating the need to monitor DCD/UCD broadcast messages which is the most time consuming step.

Based on the decoded information from MOB_NBR-ADV message, the MS becomes aware of the neighboring BSs and triggers scanning and synchronization phase. So the MS sends MOB_SCN-REQ to the SBS indicating neighboring BSs for which scanning intervals are requested. This message includes requested scanning interval duration (N frames), interleaving interval (P frames) and the number of iterations (T).

Once SBS receives MOB_SCN-REQ message, it responds with MOB_SCN-RSP message granting the request of MS following which the MS may scan target BSs (TBSs) beginning at start frame (M) during the time allocated by the SBS. SBS may also ask MS to report the scanning results by transmitting MOB_SCN-REP message.
The scanning can also be initiated by the SBS in which it sends MOB_SCN-RSP message indicating a list of recommended neighboring target BSs. IEEE 802.16e scanning procedure is shown in figure 6-13.

Figure 6-13: IEEE 802.16e Scanning Procedure
During the scanning interval, the SBS may buffer data packets addressed to MS and then transmit that data during any interleaving interval. In our work we adjust scanning interval (light scanning) and interleaving period so as to minimize packet loss. Especially when handover takes place, previous SBS transfers the buffered packets to new SBS exploiting the EPON-based distributed ring network and hence there is almost no packet drop.

One more implementation is setting pre-defined scanning threshold value (SNR value) which will be compared with MS or BS measurement to determine when the scanning procedure should start.

Depending on the value of the scanning type field indicated in MOB_SCN-REQ the MS may request either scanning only or scanning with association. The association is an optional initial ranging phase between MS and one of NBSs that may be performed during the scanning interval. IEEE 802.16e defines three levels of association:

Association Level 0 – scan/association without coordination: the target BS has no knowledge of the scanning MS and provide only contention based ranging allocations. Figure 6-14 shows association without co-ordination.

Association Level 1 – Association with coordination: the SBS coordinates the association between the MS and the requested neighboring BSs. Each NBR assigns a unique code number and transmission opportunity within the allocated region to each MS. The SBS provides pre-assigned ranging information via the MOB_SCN-RSP message.

Association Level 2 – Network assisted association reporting: In this case MS makes a list of NBSs with which it wishes to perform association. The association will then be coordinated like
Association Level 1. In this case MS is only required to send the CDMA ranging code to the NBSs and not to wait for RNG-RSP which is sent by the NBSs to the SBS over the backbone.
**Ranging and Network Re-entry**: Ranging is the process of acquiring the correct timing offset and power adjustments so that MS’s transmissions are aligned with the BS receive frame and received within appropriate reception thresholds. It consists of two procedures i.e., initial ranging and periodic ranging. Initial ranging allows a MS to join the network to acquire correct transmission parameters such as timing offset and Tx power level so that MS can communicate with the BS. Periodic ranging follows Initial ranging and adjusts the transmission parameters so that MS can maintain UL transmissions.

Ranging latency before any retry is given as:

\[
T_{\text{ranging}} = \left( \frac{2^{B_{\text{exp}}}}{N_{\text{cs}}} \right) \times T_{\text{frame}} + T_3
\]

Where;

\[B_{\text{exp}} = \text{backoff exponent}\]

\[T_{\text{frame}} = \text{frame duration}\]

\[T_3 = \text{Timeout value for receiving ranging response (50ms to 200ms)}\]

\[N_{\text{cs}} = \text{Number of slots per frame in SC PHY model or contention area}\]

\[\text{in OFDMA PHY profile}\]

The backoff exponent is characterized by the Ranging backoff start and end values. Only values between 0 and 15 are allowed. At the first transmission, the range extends over Ranging Backoff Start slots. With perceived failure the backoff range is doubled upto maximum range specified.

The Ncs attribute defines the extent of the contention area reserved for initial ranging within the frame. Minimum recommended value is 6 subchannels by 2 symbol times.
**Sharing Basic Capabilities:** After completion of ranging, MS sends SBC-REQ message to the BS informing its basic capabilities i.e., PHY and bandwidth allocation parameters. The BS responds with SBC-RSP message.

**Registration:** Registration is the process by which SS is allowed entry into the network and receives Channel Identifiers (CIDs). To register with a BS, the MS will send REG-REQ message and get responded with REG-RSP message from BS.

**Establishing Provisioned Connections:** After registration the BS will send DSA-REQ messages to the MS to set up connections for pre-provisioned service flows belonging to the MS. The MS responds with DSA-RSP message.

All these phases are shown in Fig. 6-15

Handover latency or handover delay is defined as the time interval between when the MS disconnects from the SBS until the start of transmission of the first data packet from the TBS. In HHO the total handover execution delay neglecting security consideration, $T_{\text{handover}}$ is given as:

$$T_{\text{handover}} = T_{\text{ranging}} + T_{\text{SBC}} + T_{\text{REG}} + T_{\text{DSA}}$$  \hspace{1cm} (2)

Where;

$T_{\text{ranging}} = \text{time required for MS to carry out the initial ranging process}$

$T_{\text{SBC}} = \text{time required for MS to inform on basic capabilities, SBC – REQ and SBC – RSP message exchange}$

$T_{\text{REG}} = \text{time required for MS registration with TBS,}$

$\text{REG – REQ and REG – RSP message exchange}$
Observing Fig. 6-15 and considering that these messages are exchanged sequentially between MS and BS, equation (5) can be written as:

\[ T_{\text{handover}} = T_{\text{ranging}} + 6T_{\text{Frame}} \quad \text{--------------------------(3)} \]

Where \( T_{\text{Frame}} \) is WiMAX frame size and chosen as 5 ms.

Hence equation (6) can be simplified into

\[ T_{\text{handover}} = T_{\text{ranging}} + 30\text{ms} \quad \text{--------------------------(4)} \]

Hence without association, handover delay would depend upon ranging time which can be calculated using equation (1) and can range from 5 ms up to more than 50 ms.

On the other hand handover delay in our proposed distributed ring architecture with association as explained before can skip one or several stages in the network entry process such as negotiating basic capabilities (SBC), Privacy Key Management (PKM) authentication phase, Traffic Encryption Key (TEK) establishment phase or REG-REQ message phase. Moreover ranging is performed during the scanning iterations, so all this can greatly reduce the handover delay.
Figure 6-15: The network entry process. Exchange message depend upon IEEE 802.16e standard.
The network entry messages such as SBC-REQ, SBC-RSP etc., can be exchanged using the distributed ring-based architecture saving precious handover time.

### 6.3.2 Proposed Handover Scheme

We designed a customized handover scheme and call it Optimized Distributed Hard Handover Algorithm (ODHHO). It takes into account the EPON-based distributed WiMAX network for message exchange thus offloading greatly OLT/ASN task and also saves the bandwidth on already congested backbone network. It is shown in Fig. 6-16.

The handover process begins with a decision which can be made either by the MS or the BS. A handover can be decided on many criteria e.g., when MS expected to performs better with TBS than SBS or to select best performing TBS among many NBSs. A handover decision algorithm is beyond the scope of IEEE 802.16e standard.

Once a handover is decided, it is notified through a MOB_MSHO-REQ or a MOB_BSHO-REQ indicating TBS. If MS originates handover then it sends MOB_MSHO-REQ message and will get acknowledged with MOB_BSHO-RSP. When BS initializes HO it can be mandatory or recommended. If it is a mandatory handover then MS will send MOB_HO-IND to the SBS. The MOB_HO-IND may have a HO accept or reject option. MS will send MOB_HO-IND with HO reject if it is unable to handoff to the recommended TBSs indicated in MOB_BSHO-REQ.
Figure 6-16: Optimized Distributed Hard Handover (ODH HO) Algorithm
In this algorithm, as soon as the SNR as measured by the MS from the serving BS(SBS) falls below scanning threshold, scanning interval starts which measures ‘K’ neighboring BSs(NBSs). In our work we use Association level 2 with the modification that the TBSs communicate with the SBS over the EPON-based distributed network instead of using backbone network. This not only saves the precious network resources but is also much faster. We call it Association Level 2 Distributed. Hence initial ranging also takes place during the scanning interval.

Any NBS that satisfies the following equations (5), (6) & (7), is selected as target BS (TBS) and MS sends MS-MSHO_REQ message to TBS. If MS receives MS-BHOS_RSP message from TBS then it sends MS-HO_IND to the SBS to indicate that it is handing off to the TBS. If MS handover to TBS is not successful then it sends MS-HO_IND with cancellation option to the SBS.

MS triggers handover process only on satisfying the following conditions where equation (7) is the final decision factor for the selection of TBS.

\[ \exists K \in \{ i | BS_i \text{ is a neighbor BS} \} : \]

\[ SNR(BS_k) - SNR(SBS) \geq H_1 \]

\[ Cidle(BS_k) > H_2 \times Cmax(BS_k) \]

\[ D_K = \alpha_1 \times \frac{Cidle(BS_k)}{Cidle(SBS)} + \alpha_2 \times \frac{SNR(BS_k)}{SNR(SBS)} \]

Where \( H_1 \) = SNR Threshold value

\( H_2 : 0.1 \leq H_2 \leq 1 \)

\( Cidle = BS \text{ idle(free)capacity} \)
\[ C_{\text{max}} = \text{BS maximum capacity} \]

\[ a_1, a_2 = \text{weighting factors}(\geq 1) \]

\[ D_k = \text{Decision factor for the kth BS} \]

Equation (5) is the first condition to be fulfilled by the neighboring BS for the possibility to be TBS. The main benefit of setting a threshold value is to eliminate unnecessary handovers when the SBS signal is still adequate. The fluctuations of signal strength associated with shadow fading cause a communication to be handed over back and forth repeatedly between neighboring BSs what is called the Ping-Pong effect. Selecting suitable value of the threshold value prevents so called Ping-Pong effect. Equation (6) is the second tier condition which assures that possible TBS must have specified capacity available. Equation (7) is the final decision factor for the selection for the TBS. Any BS which has the highest value of \( D_k \) is selected as TBS.

Our proposed algorithm Optimized Distributed Hard Handover (ODHHO) takes place when the network re-entry is shortened by the TBS’s previous knowledge of the MS information obtained from SBS over the distributed EPON-based WiMAX network. In this way ODHHO performs inter-BS’s communication over the local ring network. Depending on the information provided, a TBS might decide whether to skip one or several stages in the entry process such as negotiating basic capabilities (SBC), Privacy Key Management (PKM) authentication phase, Traffic Encryption Key (TEK) establishment phase or REG-REQ message phase.

It is quite obvious from the simulation results that our proposed ODHHO scheme using EPON-based ring can have a quite improvement over the typical HHO scheme as TBS gets information about the MS using the local EPON-based network, it can decide to skip SBC and REG
establishment phases. Fig. 6-17 is comparison of handover delay of typical vs. EPON-based distributed network architecture. Our proposed scheme shows much less handover delay as the network re-entry time is shortened by skipping REG, SBC etc., messages and using local network for faster communication and decision making.

![Handover Delay Comparison](image)

Figure 6-17: Handover Delay Comparison: Typical Vs. EPON-Based Distributed Architectures

### 6.4 Illustrative Examples

We developed an event driven C++ program to translate our proposed ODHHO algorithm and used in conjunction with OPNET® Simulator WiMAX module. Four different scenarios are used to stress test our proposed algorithm. The results show major improvement over typical WiMAX networks handover schemes in throughout, delay and packet loss.
6.4.1 Handover Scenario 1

One MS node is Mobile IPv4 enabled, with Home Agent (HA) set to BS_1. The MS node moves away from the Home Agent and visits 7 Foreign Agent (FA) BS nodes, before returning back to the care of the Home Agent. MS starts moving at 10 mps at around 50 seconds. Simulation setup is shown in Figure 6-18.

![Simulation Setup](image)

Figure 6-18: Simulation Setup

The first handover occurs at around 119.95 sec. Bi-directional application traffic is configured between the MS and the server node in the backbone, as follows:

- **UL**: 64 kbps mapped to the Best Effort connection.
- **DL**: 64 Kbps mapped to the Best Effort connection.

As expected the downlink traffic experiences more interruptions due to Layer 3 handover delays (via Mobile IPv4) in addition to the Layer 2 handover delays (via WiMAX MAC functionality).
The green circles correspond to areas around each BS outside which the SNR of an MS transmission to the BS drops below the scanning threshold of 27dB configured in the MS. The Mobile Station will request to enter in scanning mode (send MOB_SCN-REQ) if the downlink measurements of the serving BS falls below "Scanning Threshold". Once the smoothed SNR drops below the scanning threshold, scanning activity is started in the hope of identify other BSs as target for handover. The scanning activity statistic shows that the scanning activity ends, once the MS successfully changes its serving BS (as shown in the Serving BS ID statistic). Also each change of BS is preceded by brief initial ranging, as shown in the Initial Ranging Activity statistic and as required by the Standard during network entry.

After the MS attaches to the new serving BS, scanning still continues until the SNR at the MS receiver gets within a 'comfortable' zone (higher than the SNR threshold for scanning). As explained above, the 'comfortable' zones are represented by the green concentric circles around each BS; for example, as the MS passes by BS_5 without entering its green zone, the scanning activity continues as long as BS_5 is the serving BS.

For the sake of easy interpretation of the Serving BS ID statistics, each BS has been assigned a MAC address (or equivalently, BS ID) corresponding to its name, e.g. MAC 1 for BS_1, MAC 2 for BS 2 etc. Thus it is clear from the Serving BS ID statistic that the MS associates consecutively with BS_1, BS_2, BS_3, BS_4, BS_5, BS_6, BS_7 and BS_8, before returning to BS_1.

SNR threshold is selected to be 27 dB, scan interval (N) as 4 frames, interleaving interval (P) as 240 frames and scan iterations (T) as 10. WiMAX frame duration used is 5 ms.
Figure 6-19: WiMAX Mobility Analysis in Typical Architecture

Fig. 6-20 is obtained from Fig. 6-19 by zooming the time when MS handover from BS_1 to BS_2. To have more clarity we consider to show Initial ranging activity Scanning interval BS ID and SNR.
Similarly we get Fig. 6-21 by zooming exactly around the time when the handover takes place. For scanning interval activity -1 shows scanning is off, 0 means interleaving interval and 1 means scanning is turned on.
Figure 6-21: WiMAX handover activity zoomed in close to frame level
Figure 6-22: Throughput analysis between Typical and Distributed architectures

Figure 6-23: Average Throughput analysis between typical and distributed architectures
Figure 6-24: MS UL Throughput analysis Typical Vs. Distributed Architectures

Figure 6-25: MS DL Throughput analysis Typical Vs. Distributed Architectures
Figure 6-26: Handover Delay Comparison Typical Vs. Distributed architecture
Figure 6-27: Average Handover Analysis Typical Vs. Distributed architectures
6.4.2 More Handoff Scenarios

For the next three scenarios, some of the main configuration parameters used in the simulation are given in Table 6-1.

Table 6-2: Simulation Parameters for Handoff Scenarios

<table>
<thead>
<tr>
<th>Configuration Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame Size</td>
<td>5 ms</td>
</tr>
<tr>
<td>Frequency band</td>
<td>20 MHz</td>
</tr>
<tr>
<td>Scanning Threshold</td>
<td>35 dB</td>
</tr>
<tr>
<td>Scan Duration(N)(frames)</td>
<td>3</td>
</tr>
<tr>
<td>Interleaving Period(P)(frames)</td>
<td>255</td>
</tr>
<tr>
<td>Scan Iterations(T)</td>
<td>5</td>
</tr>
<tr>
<td>T44 Scan request retransmission timer</td>
<td>50 ms</td>
</tr>
<tr>
<td>Backoff Window Size</td>
<td>3</td>
</tr>
<tr>
<td>Ncs</td>
<td>2*6</td>
</tr>
<tr>
<td>MS speed</td>
<td>10 mps</td>
</tr>
<tr>
<td>Traffic load</td>
<td>Scenario dependent</td>
</tr>
</tbody>
</table>
6.4.3 Handover Scenario 2

In this scenario a MS handover performance is analyzed as it travels through the EPON-based WiMAX network and compared when the same MS moves across the typical WiMAX network covering almost the same distance as circumference of the ring. Simulation setup is shown in Figure 6-28. MS1 starts moving from integrated ONU1-BS1 and traverses the path across the ring. ONU1-BS1 to ONU4-BS4 set is regarded as lightly loaded, ONU5-BS5 to ONU8-BS8 medium loaded, ONU9-BS9 to ONU12-BS12 heavy loaded and ONU13-BS13 to ONU16-BS16 as super-heavy loaded. MS moves at the velocity of 10 m/s. Ring has a radius of around 1 Km and ONU-BS’s are evenly spaced. So it takes anywhere between 30-40 sec to handover to the next ONU-BS and around 650 seconds to travel complete ring as evident from Figure 3. MS1 receive data rate is selected as high enough i.e., 5 Mbps of rtsp service class which is streaming video to Video_Server attached to ONU1. 5 Mbps can account for at least 50 MS’s with data rate of 100 Kbps each. The reason for choosing one MS is to show the clarity of Handoff points and traffic received comparison with typical WiMAX model.
MS traffic received is shown in Figure 6-29 and is divided into 4 parts A, B, C and D. Part A shows the MS traffic received from lightly loaded ONU-BS’s with around 10 Mps free BS capacity, part B shows MS traffic received from medium loaded ONU-BS’s with around 5 Mps free, part C for heavy loaded ONU-BS’s with around 3 Mps and part D shows MS traffic received from super loaded ONU-BS’s with around 0.5 Mps free BS capacity. Figure 6-29 shows that as the MS1 passed through the part A and part B i.e., lightly and medium loaded ONU-BS’s respectively, both EPON-based WiMAX and typical WiMAX have similar traffic characteristics although the data rate is dropped in case of typical WiMAX when handover takes place. As the MS1 travels through heavy loaded and super heavy loaded ONU-BS’s (Parts C and D in Figure 6-29) its data rate continuously decreases respectively in case of typical
WiMAX. However as it appears in Figure 6-29., the data reception of MS remains consistent as it passes through all the regions of the ring. This is because of the interoperability of the ring and efficiently allocating the resources from lightly loaded ONU-BS’s to the ones that are heavy loaded using distributed BW allocation algorithm. Figure 6-30. shows the MS Average traffic received.

Figure 6-29: MS Traffic Received comparison

Figure 6-30: MS Average Traffic Received
6.4.4 Handover Scenario 3

In this scenario we compare the handover performance of MS’s as they traverse around EPON-based WiMAX as compare to when they move through typical WiMAX network. Each ONU is integrated with two WiMAX BS’s, so there are total 16 ONU’s and 32 integrated WiMAX BS’s. Again, ONU-BS set from 1-4 are regarded as lightly loaded, 5-8 as medium loaded, 9-12 as heavy loaded and 13-16 as super heavy loaded ONU-BS’s. Five MS’s start moving with velocity of 10 m/s from ONU1 and have trajectories that move towards ONU16 as shown in Figure 6-31. MS1 starts moving at time $t=0$ second, shortly followed by MS2, MS3, MS4 and MS5 respectively but the difference in timing is few seconds only. Each MS is transmitting streaming video which belongs to rtps service class with a data rate of 2 Mbps. Simulation setup is shown in Figure 6-31.

![Figure 6-31: Simulation Setup](image)
MS1, MS2 and MS3 receive traffic from a local video server1 attached with ONU1 whereas MS4 and MS5 receive their traffic from a video server2 beyond their local network. MS traffic received is shown in Figure 6-32 and divided into four parts A, B, C & D referring to lightly loaded, medium loaded, heavy loaded and super-heavy loaded ONU-BS’s respectively. It is evident that MS’s traffic received in case of EPON-based WiMAX configuration is pretty stable whereas traffic received decreases in the case typical WiMAX network configuration following part A. In part A i.e., within lightly loaded ONU-BS’s, MS’s traffic received is almost identical in EPON-based WiMAX and typical WiMAX network configurations. In part B i.e., medium loaded ONU-BS’s the available capacity in each of the four cells is 10 Msp (5 Msp each BS). All 5 MS’s would require 10 Mbps but as the first 4 MS’s take up 8 Mbps from both BS’s, MS5 left with 1 Mbps. In part C where available capacity is 3 Msp per BS (6 Msp for 2 BS cell), MS1 and MS2 get their fair share of 2 Mbps each henceforth the upcoming MS3 and MS4 left with 1 Mbps each whereas MS5 left with no capacity available and its data rate drops to zero as evident from figure 6-32. In part D where available capacity per BS is only around 0.4 Msp per BS, MS1 and MS2 get their share but incoming MS3, MS4 and MS5 data rate drops to zero because of scarce resources. In the case of EPON-based WiMAX network configuration, data rate is quite stable owing to efficiently allocating the resources from lightly loaded ONU-BS’s to the ones that are heavy loaded using distributed BW allocation algorithm. Figure 6-33 shows MS’s average traffic received and Figure 6-34 shows total average traffic received in both cases.
Figure 6-32: MS’s Traffic Received comparison

Figure 6-33: MS’s Average traffic Received
Figure 6-34: MS’s Total Average traffic Received
### 6.4.5 Handover Scenario 4

In this scenario we compare the upstream capacity throughput performance of EPON-based WiMAX with typical WiMAX in case of overloaded ONU-BS(s). It is assumed that two WiMAX BS’s are integrated to the ONU each with 20 MHz bandwidth. There are four heavy loaded BS’s available in the neighborhood. Upto 100 MS’s, each with data rate of 0.5 Mbps have trajectories and moving to that particular cell at random time intervals. On average 10 MS’s move into this cell every 5-10 seconds, starting at t = 30 seconds. Adaptive Modulation and coding scheme (AMCS) is used but on average 16 QAM with coding rate of \( \frac{3}{4} \) dominates. Theoretical UL capacity of two BS’s combined using 15 OFDMA symbols (350 slots) each and 16 QAM \( \frac{3}{4} \) would be around 20 Mbps whereas required throughput would be 50 Mbps as requested by 100 MS’s shown in Figure 6-35. However simulation shows, as evident from Figure 6-36, that in case of typical WiMAX throughput reaches around 30 Mbps as compared to EPON-based WiMAX where throughput reaching 50 Mbps owing to the handoff to the neighboring cells. This is because in typical WiMAX the MS is handover usually because of SNR value whereas our algorithm also looks for the capacity available in the BS and able to handoff to the neighboring cell with qualifying capacity. Moreover the distributed BW allocation algorithm in the EPON-based WiMAX network helps in load balancing.
Figure 6-35: MS’s Total Average traffic Received

Figure 6-37 shows the number of packets loss during handover and Figure 6-38 shows the average delay of EPON-based WiMAX as compared to typical WiMAX.

Figures 6-36, 6-37 and 6-38 show that not only the proposed EPON-based WiMAX network scheme has better throughput but also less packet drop and less average delay.

Figure 6-36: UL Handover Throughput EPON-based Vs. typical-WiMAX
Figure 6-37: Packet Loss analysis EPON-based Vs. Typical-WiMAX

Figure 6-38: Average Handover delay EPON-based Vs. Typical WiMAX
6.5 Conclusion

This thesis has examined the technological requirements and assessed the performance analysis and feasibility for deploying TDM-based NG-PON solutions in the access arena as multiservice, all packet-based 4G mobile backhaul RAN and/or converged fixed-mobile optical networking architecture. Specifically, this work has proposed and devised a simple and cost-effective 10G-EPON-based 4G mobile backhaul RAN architecture that efficiently transports and supports a wide range of existing and emerging fixed-mobile advanced multimedia applications and services along with the diverse quality of service (QoS), rate, and reliability requirements set by these services. The techno-economics merits of utilizing PON-based 4G RAN architecture versus that of traditional 4G (mobile WiMAX and LTE) RAN have been thoroughly examined and quantified.

To achieve the overall objective, we have utilized the existing fiber-based PON access infrastructure with novel ring-based distribution access network and wireless-enabled OLT and ONUs as the multiservice packet-based 4G mobile backhaul RAN infrastructure. Specifically, the proposed work has been carried out into two sequential phases. In the first phase, we examined and quantified the overall performance of the standalone ring-based 10G-EPON architecture via modeling and simulations. We then have assembled the basic building blocks, components, and sub-systems required to build up a proof-of-concept prototype testbed for the standalone ring-based EPON architecture. The testbed is used to verify and demonstrate the performance of the standalone architecture, specifically, in terms of power budget, scalability, and reach.

In the second phase, we have developed an integrated framework for the efficient interworking between the two wireline PON and 4G mobile access technologies, particularly, in terms of
unified NCM operations. Specifically, we have address the key technical challenges associated with tailoring a typically centralized PON-based access architecture to interwork with and support a distributed 4G RAN architecture and associated radio NCM operations.

It is shown that the overall performance of the proposed EPON-based 4G backhaul including both the RAN and Mobile Packet Core (MPC) {Evolved Packet Core (EPC) per 3GPP LTE’s standard} is significantly augmented compared to that of the typical 4G RAN, specifically, in terms of handoff capability, signaling overhead, overall network throughput and latency, and QoS support. Furthermore, the proposed architecture enables redistributing some of the intelligence and NCM operations currently centralized in the MPC platform out into the access nodes of the mobile RAN. Specifically, as this work will show, it enables offloading sizable fraction of the mobile signaling as well as actual local upstream traffic transport and processing (LTE bearers switch/set-up, retain, and tear-down and associated signaling commands from the BSs to the EPC and vice-versa) from the EPC to the access nodes (ONUs/BSs). This has a significant impact on the performance of the EPC. First, it frees up a sizable fraction of the badly needed network resources as well as processing on the overloaded centralized serving nodes (AGW) in the MPC. Second, it frees up capacity and sessions on the typically congested mobile backhaul from the BSs to the EPC and vice-versa.
7 References

7.1 Chapter 1


7.2 Chapter 2


7.3 Chapter 3


7.4 Chapter 4


7.5 Chapter 5


7.6 Chapter 6


8 Publications


