

Performance of Contention Based Access Control for a Media Frame Network

by

Teng Ge

B.Sc., Beijing University of Posts and Telecommunications, 2009

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF APPLIED SCIENCE

in the Department of Electrical and Computer Engineering

© Teng Ge, 2012
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Supervisory Committee

Performance of Contention Based Access Control for a Media Frame Network

by

Teng Ge

B.Sc., Beijing University of Posts and Telecoms, 2009

Supervisory Committee

Dr. Thomas Darcie, **Supervisor**
(Department of Electrical and Computer Engineering)

Dr. Xiaodai Dong, **Departmental Member**
(Department of Electrical and Computer Engineering)

Dr. Sudhakar Ganti, **Outside Member**
(Department of Computer Science)

ABSTRACT

Supervisory Committee

Dr. Thomas Darcie, **Supervisor**
(Department of Electrical and Computer Engineering)

Dr. Xiaodai Dong, **Departmental Member**
(Department of Electrical and Computer Engineering)

Dr. Sudhakar Ganti, **Outside Member**
(Department of Computer Science)

The idea of a Media Frame network (MFN) was proposed very recently for solving the explosively growing demand for end-to-end large file transfers. This networking method combines the advantage of high transmission speed from optical networks and flexibility and fast header parsing from electronic networks. The MFN is based on very large data units or media frames (MF) compared to IP packets. Due to the logical continuity, transporting data in a media frame network largely reduces the power consumption in the intermediate nodes and routers. Currently the backbone of media frame network has been studied. The remaining challenge is to devise a system solving the problem of transporting MFs through access networks (i.e., the last mile) connecting customers to the backbone networks. To our knowledge, no other research activity regarding this challenge has been reported. If this challenge is overcome and if the overall concept is accepted, the MFN could be a very important step in the evolution of the Internet.

This thesis focuses mainly on the access network. For the first time, a solution is proposed to establish the ability to transport media frames over a standard PON (e.g. Passive Optical Network) architecture. Because of the unique properties of the media frame network, the physical layer model and transport protocols must be rebuilt. Referring to the ITU-T G.987 recommendations, the physical layer is built based on the XGPON specification. In this thesis, the initialization protocols, bandwidth allocation plan, OLT-ONU (OLT: Optical Line Terminal, e.g. central office. ONU: Optical Network Unit, e.g. customer side box) negotiation protocols are designed. Different schemes for each protocol are proposed, with simulation support based on Omnet++. For the transmission of a 7GB file on average, different transparency degrees under different traffic conditions are compared, and the tradeoffs among essential factors are investigated.

Table of Contents

Supervisory Committee	ii
ABSTRACT	iii
Table of Contents	iv
List of Tables	vi
List of Figures.....	vii
List of Acronyms.....	viii
Acknowledgements.....	ix
Dedication	x
Chapter 1 Introduction	1
1.1 Background	1
1.2 Media Frame Network Introduction	3
1.3 Passive Optical Network Introduction	4
1.3.1 PON Fundamentals.....	4
1.3.2 Important Protocols of PON	5
1.3.3 XGPON.....	8
1.4 Access solution for Media Frame Network	10
1.5 Thesis Outline	11
Chapter 2 Media Frame Networking.....	12
2.1 Network Topology.....	12
2.1.1 Hierarchical Topology.....	12
2.1.2 Media Frame Network Topology	14
2.2 Media Frame and Media Chain.....	15
2.3 Networking Methodology	16
2.3.1 Signaling and Control	16
2.3.2 Media Chain Scheduling.....	16
2.3.3 Synchronization.....	18
2.3.4 Directly Routable Frames (DRF)	18
2.4 Benefits and Advantage of MFN	18
Chapter 3 Fundamentals of Access Network Design for MFN.....	21
3.1 Challenges and Choices.....	21
3.2 Physical Layer Design	21
3.3 Logical Structure Design	25
3.3.1 ONU Design.....	25
3.3.2 OLT Design.....	26
3.4 Conclusion.....	26
Chapter 4 Protocol Design	28
4.1 Initializing Procedures	28
4.2 Channel Sharing Schemes.....	28
4.2.1 Fixed Transparency Degrees	29
4.2.2 Variable Transparency Degrees	29
4.3 Negotiation Protocols.....	30
4.3.1 Request Resending Based Negotiation	31

4.3.2 Queue Based Request Processing	32
4.3.3 Analysis on Time-Cost Comparison	33
4.3.4 Downstream Traffic	34
4.4 Conclusion.....	35
Chapter 5 System Performance Analysis and Simulations.....	36
5.1 System Performance Indicators	36
5.2 Simulation System.....	37
5.3 Simulation for Fixed Transparency Degree.....	38
5.3.1 TD of 1 Media Chains.....	38
5.3.2 Larger fixed TD Simulation.....	40
5.4 Analysis for Variable TDs.....	44
5.4.1 Combination of TD1 and TD2	44
5.4.2 Max TD of four	44
5.5 Summary	46
Chapter 6 Conclusion.....	47
Contributions of This Thesis.....	47
Future Work	48
Bibliography.....	49

List of Tables

Table 3-1 Candidate Wavelength Bands and Potential Problems.....	22
Table 4-1 Cases to Compare.....	33

List of Figures

Figure 1.1 IP Traffic Monthly Volume Forecast [2]	1
Figure 1.2 Forecast of monthly consumer Internet traffic[2]	2
Figure 1.3 Number of transferred files vs. file size [3].....	2
Figure 1.4 Block diagram of router showing fractional energy consumed in different blocks [6].....	3
Figure 1.5 Typical GPON Physical Architecture [10].....	5
Figure 1.6 Ranging Procedure of PON [11].....	6
Figure 1.7 Downstream Data Traffic in TDM PON.....	7
Figure 1.8 Upstream Data Traffic in TDM PON [12]	7
Figure 1.9 Typical DBA Procedure [11]	8
Figure 1.10 Co-existence of GPON and XGPON solution [15]	9
Figure 1.11 Wavelength Candidates for XGPON [16].....	10
Figure 2.1 Typical Hierarchical Network [5].....	13
Figure 2.2 Media Frame Overlay Network [5].....	14
Figure 2.3 Two MCs, DRFs and Voids in a Channel [5].....	15
Figure 2.4 Path Reservation and Timing.....	17
Figure 3.1 MF-PON Splitter Architecture Options [15]	23
Figure 3.2 MFN Access Network Overlay upon XGPON System.....	24
Figure 3.3 WDM on MF-PON.....	24
Figure 3.4 Original GPON ONU Structure	25
Figure 3.5 OLT Functional Blocks.....	26
Figure 4.1 Illustration of Fixed Transparency Degree of four	29
Figure 4.2 Dynamic Transparency Degrees	30
Figure 4.3 One combination of maximum numbers of TD.....	30
Figure 4.4 Request resending based negotiation flow chart	31
Figure 4.5 Queue based request processing flow chart	33
Figure 5.1 Access network topology for simulation.....	37
Figure 5.2 Media chains with TD of one.....	38
Figure 5.3 Total time and in-queue time for TD1	39
Figure 5.4 Channel utilization and blocking probability for TD1.....	40
Figure 5.5 Total time and in-queue time for TD2	41
Figure 5.6 Channel utilization and blocking probability for TD2.....	41
Figure 5.7 Results for fixed TD4	42
Figure 5.8 Queue length comparison between fixed TDs.....	43
Figure 5.9 Time values for fixed TD of eight.....	43
Figure 5.10 System performance of TD1 and TD2 combination.....	45
Figure 5.11 Total time comparison between TD123 and TD24	45
Figure 5.12 Utilization comparison between TD124 and TD24	46

List of Acronyms

CAD	Channel Availability Degree
CP	Control Packet
DRF	Directly Routed Frame
FIFO	First In First Out
GPON	Gigabyte Passive Optical Network
MAB	Media Access Bypass
MAR	Media Aggregation Router
MBR	Media Backbone Router
MF	Media Frame
MFN	Media Frame Network
MC	Media Chain
ODN	Optical Distributed Network
OLT	Optical Line Terminal
ONU	Optical Network Unit
PMD	Physical Media Dependent
PON	Passive Optical Network
RTD	Round Trip Time
TD	Transparency Degree
WDM	Wavelength Division Multiplexing
XGPON	10-Gigabyte Passive Optical Network

Acknowledgements

My time as a graduate student at UVic has been supported by many kind people. This page, undoubtedly incomplete, attempts to recognize those who have been helpful along the way.

First of all, I would like to thank my graduate supervisor, Dr. Thomas Darcie. His support, guidance and encouragement have helped me so much throughout my graduate studies. I also thank Dr. Sudhakar Ganti for his great help on the simulation setup.

I also want to thank all the people in my group. I am grateful to Ilijc Albanese, without whom I would not finish my research project and this thesis. Many thanks to Rejoy Isaac, who introduced me to UVic and my Lab with a warm heart before I came to Canada. Special thanks to Dr. Jinye Zhang (James), Dr. Min Qiao (John), Yuanjie Pang, Barmak Heshmat and Hamid Pahlevaninezhad, who helped me immensely along these years.

Outside the lab, I am grateful to all my friends, with whom I spent many happy times. I thank Zhuangzhuang Tian, who gave me the most generous help when I arrived Canada. I also thank my dear friends, Yang Song, Chenyuan Wang, Ping Li, Dan Li, Moyuan Chen, Xi Tu, Jun Zhu, Bojiang Ma, Yingduo Chen and Congzhi Liu, without whom I would not have had such a wonderful time. Many thanks to my soccer buddies, Jin Zhang, Yunqian Luo, Zhe Yang, Cheng Chen, Yimian Du, Ming Lu, Dong Zhang. With them I had my most enjoyable soccer games and shared many exciting moments.

Finally, I thank Lan, who deserves more acknowledgements than that can fit on this page. Her unquestioning support and uncanny wisdom have been invaluable through the ups and downs of this journey.

Dedication

*To my dearest parents, Xinsheng and Jun
and my beloved Lan*

Chapter 1

Introduction

This work focuses on the design of an access solution for a media frame network. A media frame network (MFN) is a very promising new network architecture, designed specially for the very large file transfers and lower power consumption in routers. An access network for MFN is built on the infrastructure of a standard passive optical network, which includes OLT (optical line terminal, i.e. the central office component of PON), ODN (optical distribution network, i.e. the optical fibers), and ONU (optical network unit, i.e. the customers side of the PON). It provides a solution for the MFN access network managing bandwidth allocation, conflict handling and task scheduling. This chapter presents an overview of the concepts of access networks, passive optical networks and media frame networks.

1.1 Background

Internet traffic data volume has been increasing explosively year by year. The annual IP traffic data will reach 28 EB (1 EB = 10^9 GB) by the end of 2011, and as expected, the worldwide Internet traffic will enter the ZB (1 ZB = 1000 EB) range by the year of 2015[1].

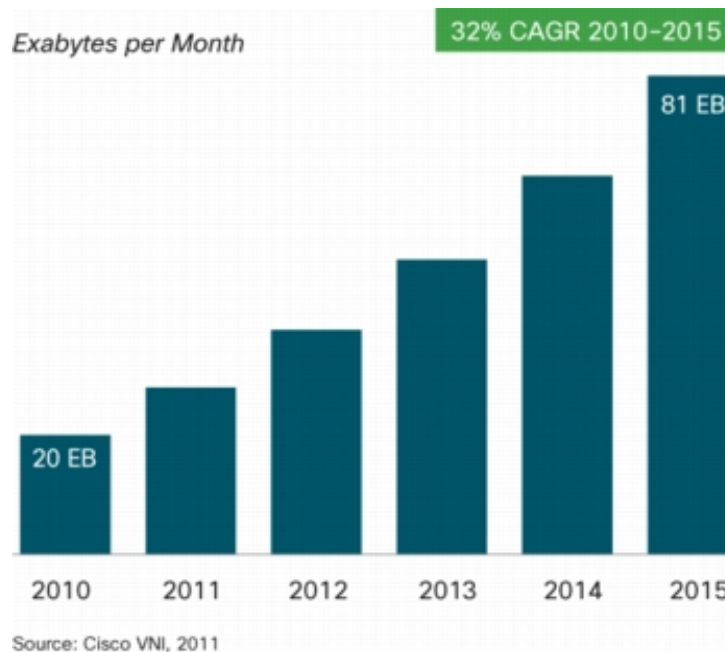


Figure 1.1 IP Traffic Monthly Volume Forecast [2]

As figure 1.1 shows, the aggregate monthly IP traffic volume is expected to grow to 81 EB by 2015, up from 20EB per month in 2010. The main reason for this unprecedented increase is Internet video being popular. From the forecast numbers in figure 1.2, Internet

video will reach 50 percent of consumer Internet traffic, and in 2015, the video traffic volume will be over 30 EB per month.

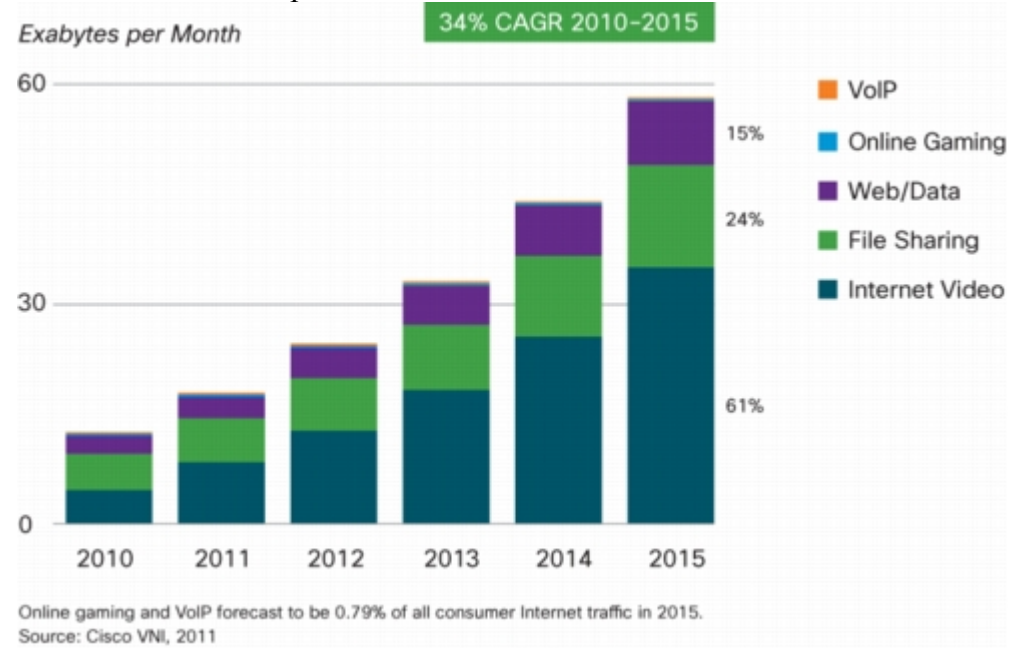


Figure 1.2 Forecast of monthly consumer Internet traffic[2]

Another recent study [3] supports that transmitting and sharing large size files that are becoming the dominant components of bandwidth transacted, and the number of transferred files decreases as file size increases, which is illustrated in figure 1.3. The redundant parts indicate that a significant portion of the files transferred is the same.

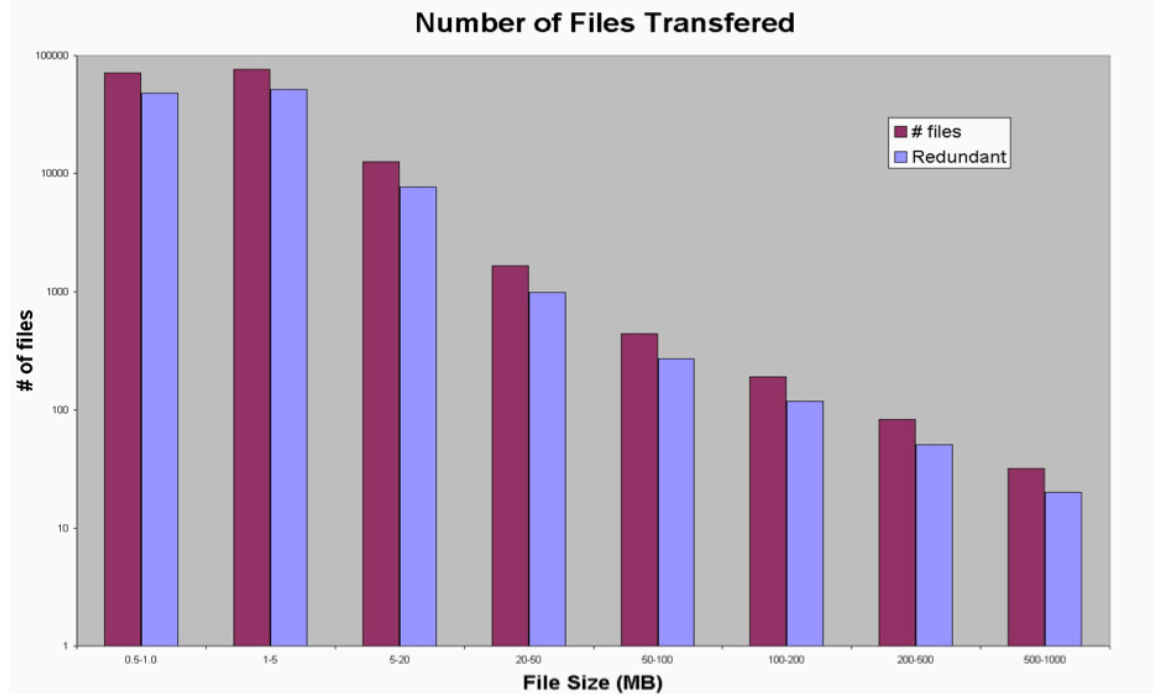


Figure 1.3 Number of transferred files vs. file size [3]

Given this growth, power consumption has become a significant issue in building data centers and server farms [4]. Figure 1.4 clearly shows the fractional energy consumed in different functional parts of IP routers. The biggest energy consumers in a router are the forwarding engine and power supply inefficiency. Since the energy consumed in the forwarding engine is driven by the number of processed packets, it would help to enhance the power efficiency by transferring fewer packets with larger size [5]. In addition, less heat will be generated due to the power efficiency improvement, and power wasted by the fans and blowers also decreases.

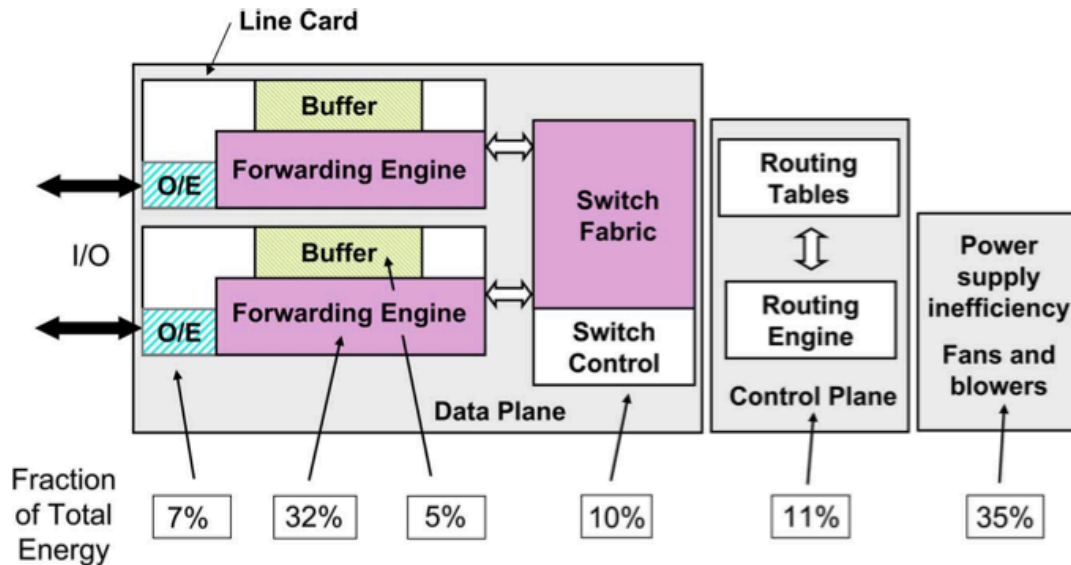


Figure 1.4 Block diagram of router showing fractional energy consumed in different blocks [6]

Under this background, a new method of networking has been proposed—the media frame network (MFN). It operates on much larger data packets, the media frames, to facilitate large file transfers, and lower the energy consumption in intermediate routers.

The core MFN is still in the design stage, and the major tasks of this thesis are designing effective ways to traverse the last-mile in this network, and provide some comparisons between different methods.

1.2 Media Frame Network Introduction

To meet the challenges brought by increasing demands of high-speed transmission and sharing of large files, researchers have proposed a few solutions in recent years. They all tried to break the constraints in router design, which is the definition of Ethernet packets containing up to 1500 bytes. Optical burst switching [7] and optical flow switching [8] provide up to an entire wavelength for gigabyte-scale file transmission, but these optical approaches are not embraced by industry [5]. Also, the power efficiency of optical

switching is subject of debate [6]. Study [9] shows that optical burst switching does improve the throughput capacity and transmission efficiency, but the bottleneck in control plane would bring limitation to large burst size and place a lower bound on the holding times of optical channels.

A media frame network combines the advantages of electronic switching, control and buffering, and optical DWDM transmission. It creates an overlay network optimized for sharing and distributing very large files. The formatted unit of data is defined as media frame (MF). Each media frame is stuffed with typically 10 megabytes of data, and a large media file fills many such media frames. For very large files, these media frames will be concatenated as a media chain (MC) and routed in an end-to-end approach through the intermediate network nodes. The media chains are considered as indivisible entities through the entire network, which minimizes header computation and buffering with efficient large volume data transport. In addition, it would be inefficient for this end-to-end admission control when individual media frames are scheduled. To avoid this condition, directly routed frames (DRF) are introduced for the contention-based connectionless routing of individual MFs, placed in the interstitials between MFs in a media chain [5]. Hence, MCs are routed swiftly following end-to-end set-up, while the DRFs are processed more conventionally.

1.3 Passive Optical Network Introduction

Passive optical network (PON) is the most important class of fiber access system in the world today [10]. It offers longer transmission distance from the central office to users' premises, and minimizes the fiber deployment in the infrastructure. Due to the deep fiber penetration, end users obtain much more bandwidth compared to conventional access systems.

1.3.1 PON Fundamentals

PON technologies have emerged since late 80s, but in the last few years, the commercial standards have matured and vendors have started to implement this technology. Current existing PON standards include APON, BPON, EPON and GPON. APON/BPON are the earliest standards proposed by ITU-T study group 15. APON (ATM-PON) and BPON (Broadband PON) are different aliases of the TDM-PON system based on ITU-T G.983 series standards [11]. EPON stands for Ethernet PON, is based on the IEEE 802.3 standard exclusively for Ethernet and IP traffic, and it is widely implemented in Asia and Europe. GPON (Gigabit PON) is a capacity enhanced standard derived from APON/BPON, as defined by ITU-T recommendation series G.984. GPON can transport not only Ethernet, but also ATM and TDM traffic by using a GPON encapsulating method (GEM). Nowadays GPON is widely deployed in North America for commercial use, and the capacity has reached 2.5 Gb/s downstream and 1.24 Gb/s upstream. In recent years, the concept of 10G-PON (XGPON) has become popular and ITU is establishing commercial standards. Its capacity for both uplink and downlink could reach 10Gb/s at an

advanced phase. Currently, the G.987 recommendations define the basic specifications for physical layer.

The architecture of a PON system consists of an Optical Line Terminal (OLT), Optical Distributed Network (ODN) and Optical Network Units (ONU). Figure 1.5 represents the typical components of a PON system. Starting at the OLT, which is located in a central office, only one single mode fiber runs to the optical power splitter. At this point, the optical signal is simply divided by the splitter into a number of separate paths to end-users. All the fibers and optical splitter device are referred to as ODN. The number of separate paths is defined as splitting ratio, varying from 2 to 128, considered as an important parameter of PON. The ONU is located either in a user's premises or at the curb, converting the optical signals from the fiber to electric signals for the subscribers. On the other side, the OLT connects the PON and WAN (Wide Area Network), controlling all the traffic in and out. Commercial vendors may also add video overlay or voice overlay upon the PON by adding an extra wavelength, to enhance the utility efficiency of the infrastructure. This is known as double play or triple play.

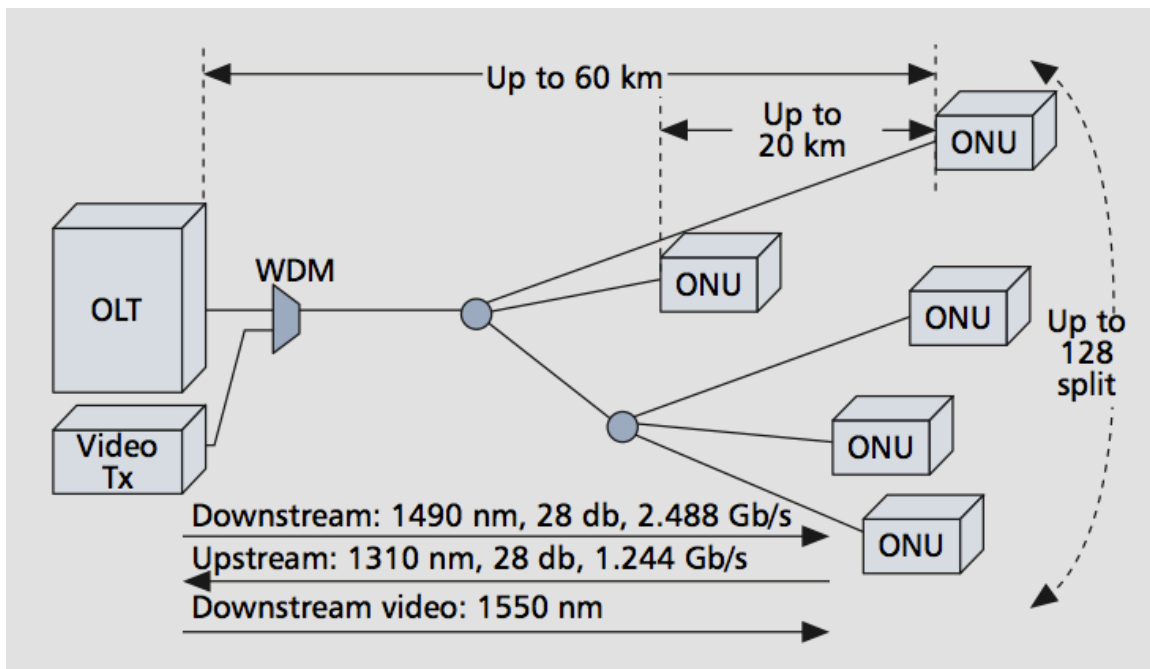


Figure 1.5 Typical GPON Physical Architecture [10]

1.3.2 Important Protocols of PON

The entire PON system consists of large numbers of ONUs, each of which has varying physical and logical characters. To organize all of the ONUs as a functional unit, some principles must be implemented.

- Ranging

Since ONUs can lie at varying distance from the OLT, the transmission delay from each ONU is unique. Therefore, it would be hard to schedule traffic and synchronize transmission time without an appropriate adjustment. The motivation of the ranging protocol is to equalize the logical distance between OLT and all ONUs, in spite of the fact that they are at different ranges spatially.

In the ranging procedures, the transmission time is scaled by RTD (round trip delay). The steps are as follows:

1. In the initializing phase of a PON system, the OLT transmits a signal with a specified ONU serial number, to measure the *RTD* of this certain ONU.
2. Upon receiving this request signal, the ONU that matches the serial number in the request signal sends a responding signal back to the OLT, while the other ONUs ignore this request.
3. As receiving the responding signal, OLT measures the time delay as the *RTD* after confirming the serial number.
4. OLT notifies an *Equalization Delay* ($=Teqd - RTD$) to the ONU. Here the *Teqd* is a constant number and the maximum *RTD* value in the PON system [11]. For instance, in the case of 30km maximum distance, the *Teqd* is approximately 0.3ms (presuming the refractive index of fiber is 1.5). Figure 1.6 illustrates this procedure in time line.

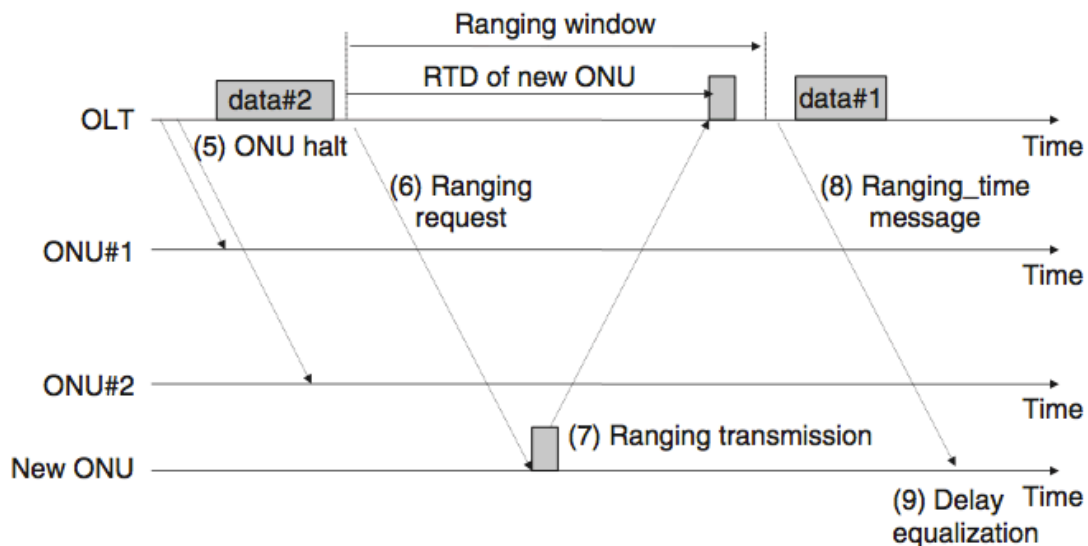


Figure 1.6 Ranging Procedure of PON [11]

5. The ONU memorizes this equalization delay it receives, and delays subsequent upstream transmission by this value.

Through this ranging process, all ranged ONU become equivalent on RTD that is logically same distance from OLT.

- Dynamic Bandwidth Allocation

In the downstream direction, the data traffic transmitted from the OLT passes through a $1 \times N$ passive power splitter and reaches each ONU. Because of the broadcasting nature, all the traffic in the downstream direction shares the same channel. The data are selectively extracted by the destination ONU (Figure 1.7).

In the upstream direction, the data sent from each ONU, through the passive combiner, reaches the OLT as the only destination. A PON needs to implement some arbitration mechanism to avoid collisions, since data from different ONU transmitted simultaneously may collide at the combining point. The PON channel could be shared in time-domain (TDM PON) or in frequency-domain (WDM PON). Considering the equipment and maintenance costs, a TDM sharing mechanism is widely implemented on current commercial PONs. Figure 1.8 demonstrates the typical upstream in TDM PON.

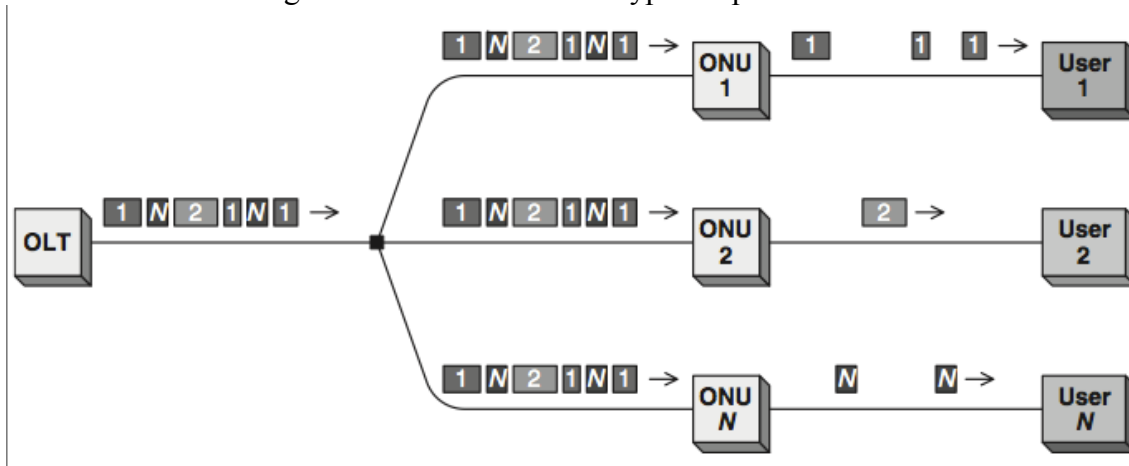


Figure 1.7 Downstream Data Traffic in TDM PON

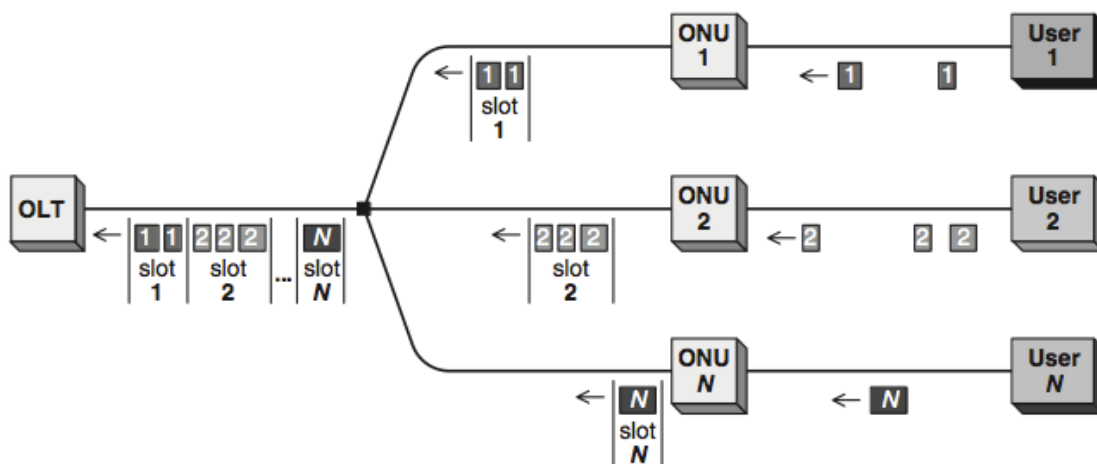


Figure 1.8 Upstream Data Traffic in TDM PON [12]

Upstream bandwidth allocation can be divided into two methods, fixed bandwidth allocation (FBA) and dynamic bandwidth allocation (DBA). Considering the high efficiency of bandwidth utilization, DBA is embraced in most cases. Figure 1.9 presents a typical explanation on how DBA operates.

1. The ONU stores the upstream traffic received from the user in the buffer.
 2. The data volume stored in the buffer is transmitted to the OLT as a request at a time prescribed by the OLT.
 3. The OLT specifies the transmission start time and available duration to the ONU as a grant, taking into account the notified volume and service specifications.
 4. The ONU waits for the granted time and then transmits the OLT specified data volume.
- [11]

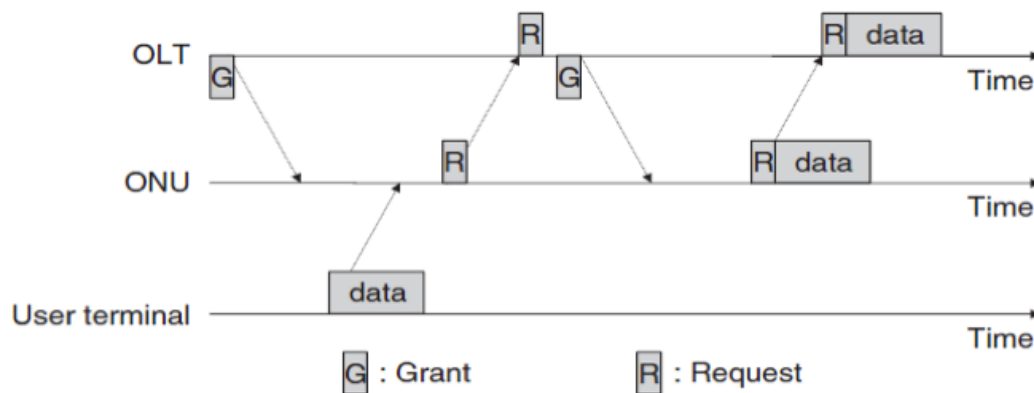


Figure 1.9 Typical DBA Procedure [11]

1.3.3 XGPON

In 2006 the FSAN (Full Service Access Network) group began to contemplate a new standard for commercial PON [13], as a successor of GPON able to transmit beyond 10Gb per second. After a few years of research and discussion, the ITU-T G987 was recommended, defining the commercial standard of XGPON, finalized in October 2010. The world's first XGPON field trial was conducted by Verizon in Taunton, MA, in late 2009 [14], and proved the XGPON has fully compatible with current GPON infrastructure, but supporting 10Gbps line rate.

To assure the smooth migration from legacy PON to NGPON, XGPON is defined as two phases. Phase 1 supports 10 Gbps upstream and 2.5 Gbps downstream while phase 2 is capable of 10 Gbps for both uplink and downlink. The current solution to provide XGPON service over GPON architecture is achieved by adding an XGPON overlay through WDM technology. In order to get simultaneous XGPON and GPON working, a WDM1r combiner/splitter is implemented in the network, as illustrated in Figure 1.10 [15].

The wavelength channels allocated for XGPON are selected from several candidates (Figure 1.11). Considering the coexistence with GPON and the high capacity of line

speed, two separate wavelength channels are required for downstream and upstream. For downstream, the wavelength band of 1575-1580 nm seems the only option left, and the fiber window is limited to 1580 nm [16]. Besides, this band has an additional advantage as it is already selected as downstream channel for 802.3av draft standard (10G-EPON).

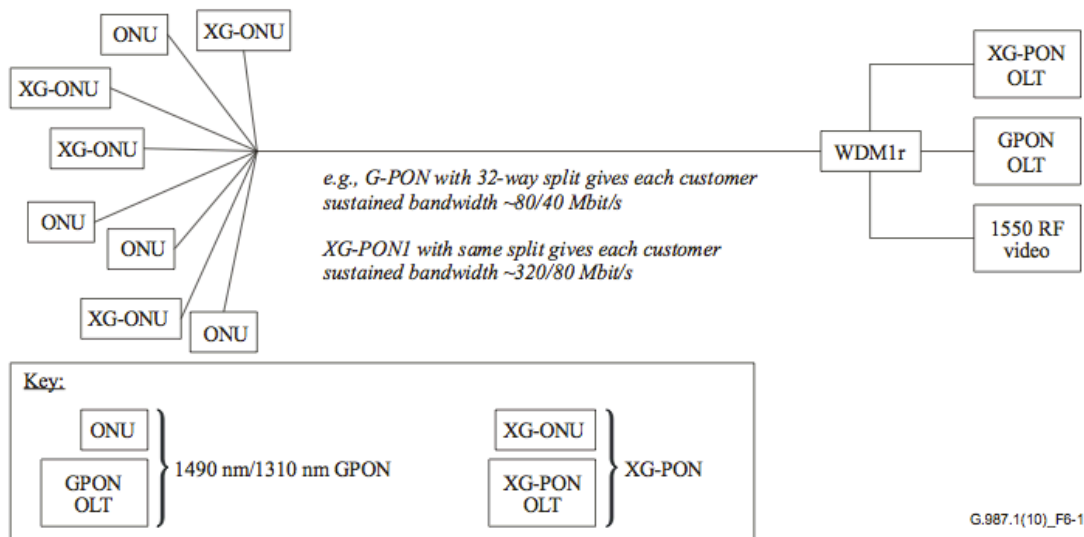


Figure 1.10 Co-existence of GPON and XGPON solution [15]

For the upstream, there were a few options to consider. Five candidate channels are available. However, after meticulous evaluation, only option E (1260nm - 1280nm) was selected. Channel A (1590nm - 1610nm) was dropped because this band may be not fully compatible with passive optical components. As many significant PON systems and video broadcasting use the 1540-1560 nm band, option B (1530nm - 1560nm) was rejected. Because the current GPON ONU may not be sufficiently isolated from this wavelength, and the ONU at such a wavelength may be costly, candidate C (1530nm - 1540nm) was eliminated as well. Besides, due to the fiber loss characteristic, it is quite difficult to achieve low loss in this channel, therefore channel D (1340nm - 1360nm) was excluded [13].

Although the XGPON was supposed to share the same ODN as GPON, and the G984 standard clearly specifies what the characteristics are: 28dB of loss in 1260-1360 nm band and 1480-1580 band [17], the power budget of XGPON cannot be reused directly from GPON because of two reasons: the utilization of the WDM1r filter and the inevitable deviation due to practical deployment [13]. Given these factors, the FSAN group finally selected two "nominal" power budgets, which are nominal 1 (29dB) and nominal 2 (31 dB) [18]. In addition, the coexistence of GPON and XGPON on an ODN drives the extended power budget requirement, allowing for an additional split in the ODN with appropriate margins, or alternatively an increase in the supported system reach.

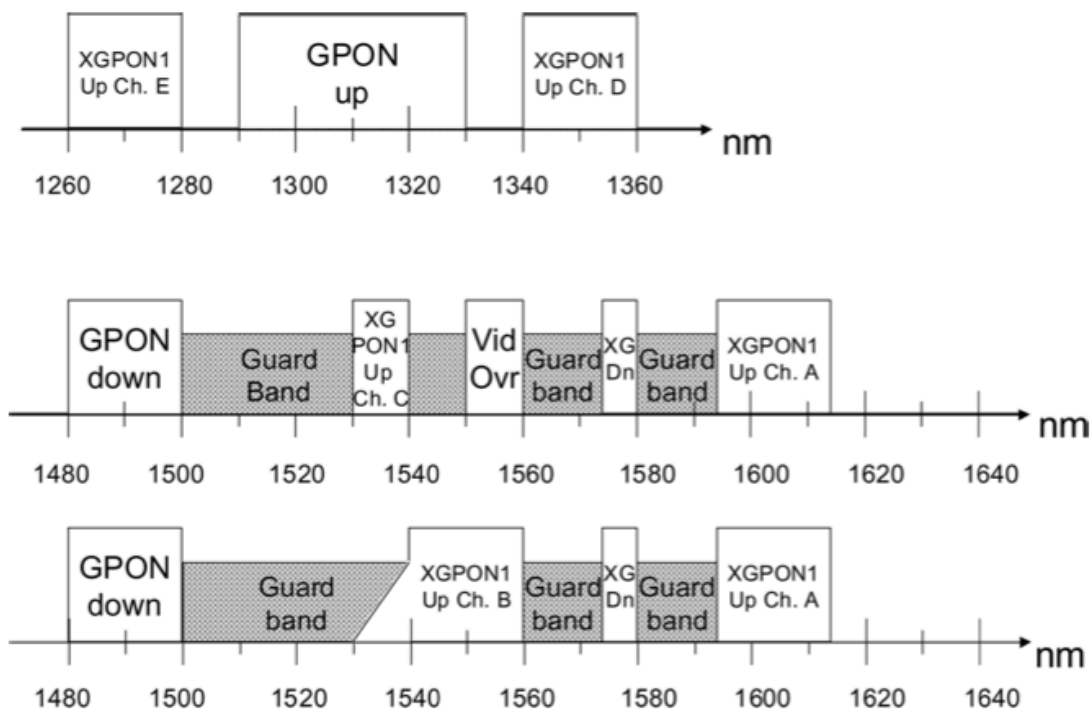


Figure 1.11 Wavelength Candidates for XGPON [16]

1.4 Access solution for Media Frame Network

The access network for the MFN must provide high-speed media frame transmission for the user uploading and downloading. The data rate in the core network is typically designed as 40 Gbps, and to assure the performance, the access network should be able to provide at least 10 Gbps for both directions. After reviewing the existing access solutions, the XGPON phase 2 could be a qualified candidate.

However, these access solution standards are all designed for the IP network, which is based on very dynamic IP packet switching. The media frame network operates upon the media frame, which is hundreds of times larger than IP packets. To accommodate media frame transfer, a special protocol layer must be redesigned above the XGPON physical media dependent layer.

The access network layer model can be simplified to three layers: physical, protocol and application layer. At this stage, the design is mainly focused on protocol layer, since the physical layer is specified in ITU-T Recommendations, and the application layer is related to some future higher-level application.

The challenge to fit the media frames in XGPON is focused on the channel scheduling. Modern dynamic channel scheduling methods are designed for individual IP packets, which means every packet could be buffered or delayed at any intermediate router

randomly, without corrupting the file reception. In the MFN, this is the case for DRF (Directly Routed Frame), which can be considered for scheduling purposes as very large individual packets. However, large files are transmitted as orderly concatenations of MFs into MCs. Media frames inside the MCs cannot be delayed since this will break the chaining sequence.

1.5 Thesis Outline

The rest of this thesis is organized as follows. Chapter 2 presents the details of how a media frame network works, as well as the configuration of media frames. The advantage and benefit brought by media frame network will also be investigated. In chapter 3, fundamentals of the access networks are studied, concentrating on the component structure and physical layer design. The protocol design of the access network is presented in chapter 4. Since there are strong differences between MFN and conventional IP networks, data transmission protocols must be redesigned to accommodate the media frame transport. Chapter 5 presents the access network performance, and provides numerical analysis based on the simulation results. In chapter 6, the contributions of this thesis will be summarized, conclusion drawn and future work discussed.

Chapter 2

Media Frame Networking

In this chapter, the working principles of a media frame network will be presented. Comparing to a conventional network topology, the MFN network architecture introduces some new components as an overlay to facilitate the media frame transmission. These include are MCI (media client), MCa (media cache), MBR (media backbone router), MAR (media aggregation router) and MAB (media access bypass).

2.1 Network Topology

The network topology is the layout pattern of the interconnections of a network. It refers to how data is transferred in a network as opposed to its physical design.

2.1.1 Hierarchical Topology

Synchronous Digital Hierarchy (SDH) and Synchronous Optical Network (SONET) are standardized multiplexing protocols that transfer multiple digital bit streams over optical fiber using lasers. They represent the typical hierarchical network topology and are deployed worldwide. The hierarchical topology consists of core network, metro area network and access network.

The core network plays the central part of the entire network, with huge capacity, connecting all of the sub-networks together and providing paths for the exchange of information between them. At the nodes of the core network, optical add-drop multiplexers interconnect with metro-area networks. The core network usually has a mesh topology that provides any-to-any connections among devices on the network. The devices and facilities in the core network are ADM switches and routers.

As a sub-net of the core network, the metro-area network (MAN) spans a city or a large campus [19], connecting the core network through core router and edge router. Typically the MAN interconnects large numbers of access networks, with high capacity backbone technology. The access networks communicate the data traffic directly to the service subscriber.

This hierarchical network, shown in figure 2.1, represents the current state-of-art comprising various sizes of routers (core, edge and access) connected to a transport network through various add-drop multiplexers. In present implementations the add-drop multiplexers would use SONET, but this reference architecture also supports optical add-drop technology. The interconnections between various carriers through peering

arguments, typically through regional or metro-area exchange points, are very important components.

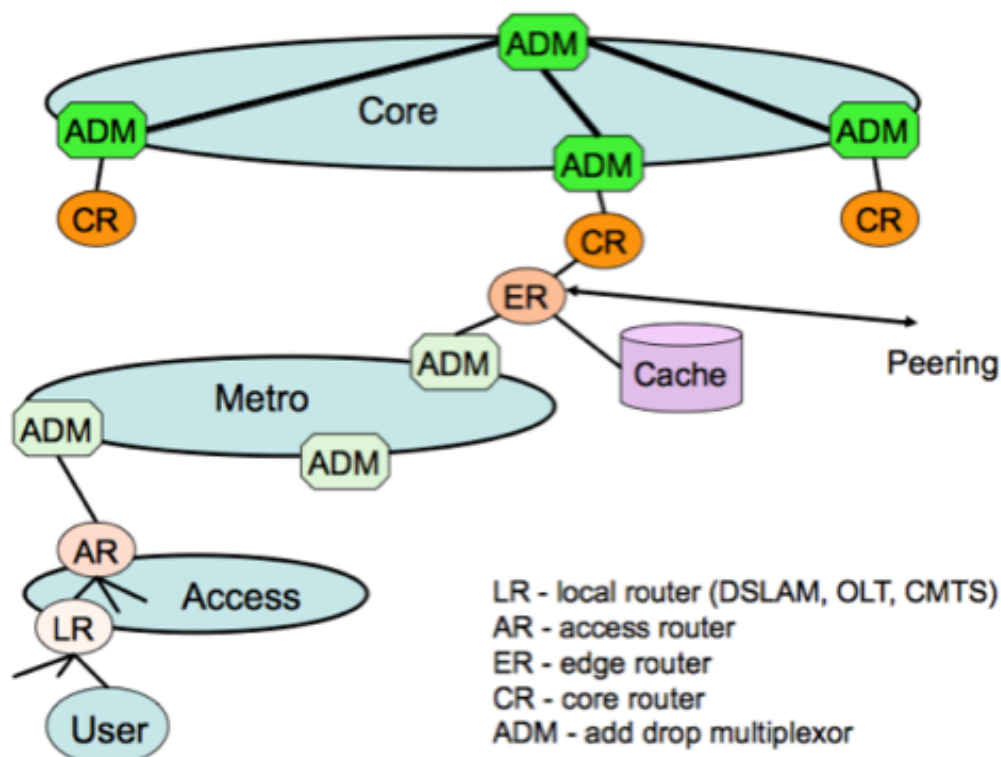


Figure 2.1 Typical Hierarchical Network [5]

Typical examples of transporting large files that are widely used today can be superimposed onto this reference architecture. These examples include: [5]

- 1) Downloading from regional cache: A most common example is a content delivery network (CDN). Large amounts of media files are stored in regional caches, which are usually distributed near the edge routers, and delivered through the MAN and access network to end users. These downloads contribute a significant portion of traffic growth.
- 2) Distributing media files from source to cache: This is to facilitate the content delivery network. Usually the media files are transmitted across the core network to distribute in the regional cache-it is mostly conducted by the CDN service provider.
- 3) End-to-end file transfer: This is also known as peer-to-peer (p2p) transfer. The p2p applications provide a very important fraction of bandwidth utilization, particularly in Metro and access network.
- 4) Transactions between different carrier networks.

2.1.2 Media Frame Network Topology

In order to operate the media frame network successfully, the media frame chains must be transferred through the hierarchical network efficiently. A feasible solution, still in research, is overlaying a media frame layer upon the reference architecture, depicted in figure 2.2.

Each network node parallels a traditional counterpart to support media frames and to co-exist with the current infrastructure. At the end user side, MCI (media client interface) supports generating and terminating the media frames. MAB (media aggregation bypass), MAR (media aggregation router) and MBR (media backbone router) connect the different hierarchical networks. Media cache (MCa) is regional storage cache at the metro-area network.

At the access network, the main task is unloading the media frame chains from a higher speed network and transferring them to users premises. Two approaches were proposed at the design stage. One is a more conservative way-the media frames and media chains are assembled and disassembled at MAR, involving an application operable between MAR and MCI such that only the existing broadband access networks are used. The other method is more aggressive and also the concentration of this thesis, involving upgraded PON technology with WDM overlay. The MF and MC are then assembled at user terminal directly.

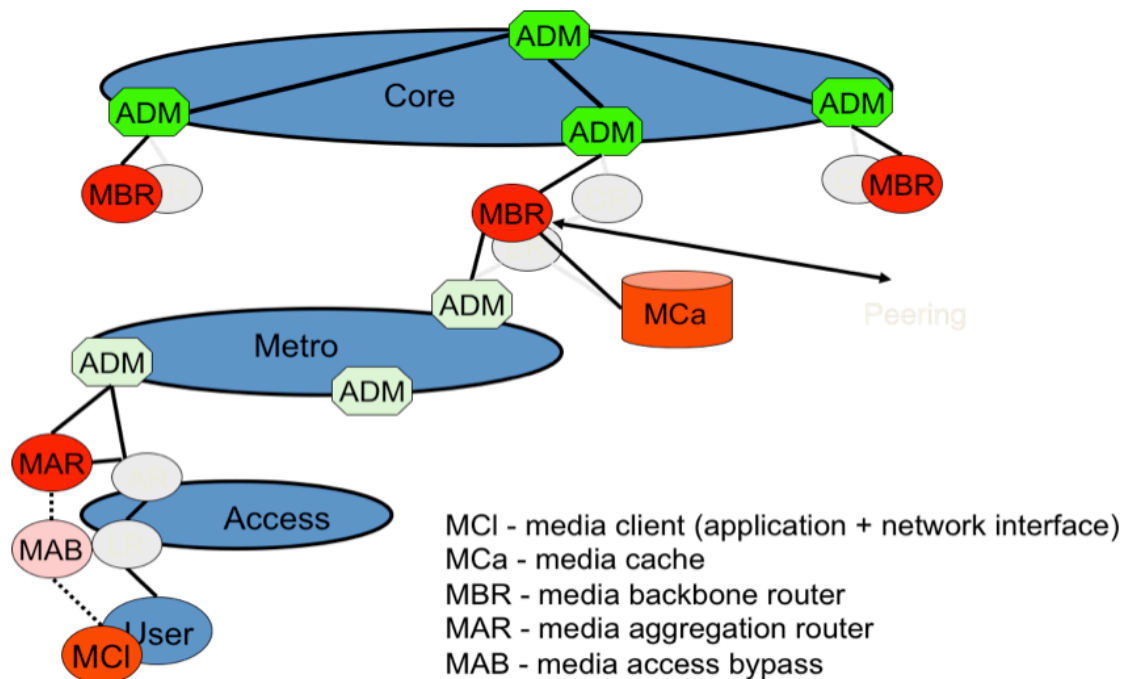


Figure 2.2 Media Frame Overlay Network [5]

The media frame network does not require specially designed physical layer provisions and restoration to support MFs. The transport network is designed to be compatible with

existing SONET and Ethernet, due to the successful convergence between SONET OC192c and 10GE WAN standards [5].

2.2 Media Frame and Media Chain

A Media Frame is introduced as the basic data unit in a media frame network, which is mentioned earlier in chapter 1. The media frame works as data container structure, stuffed with typically 10MB of payload plus an overhead. The overhead includes all the control information, such as source/destination address, media chain information, priority, coding and other management information.

Typically, when transferring a file that is much larger than 10MB, the file will be divided into many MFs, which are then concatenated into a media chain (MC). Each MC is assembled by a certain manner, e.g. transparency degree, defined as the number of interstices between two MFs plus one (considered as “period”). Since the overhead in each media frame can contain the same information about the media chain, only one header is required to individuate and identify the whole chain.

Media chains are illustrated in figure 2.3. Transparency brings large flexibility to media chain. Without transparency, concatenations of a very large file would introduce substantial latency by occupying the bandwidth resource for considerable duration (1MF = 10MB ~ 2ms @ 40Gbps OC768 transmission). In addition, the transparency also simplifies routing and scheduling for media chains, allowing them pass through intermediate nodes directly and swiftly due to the logical continuity.

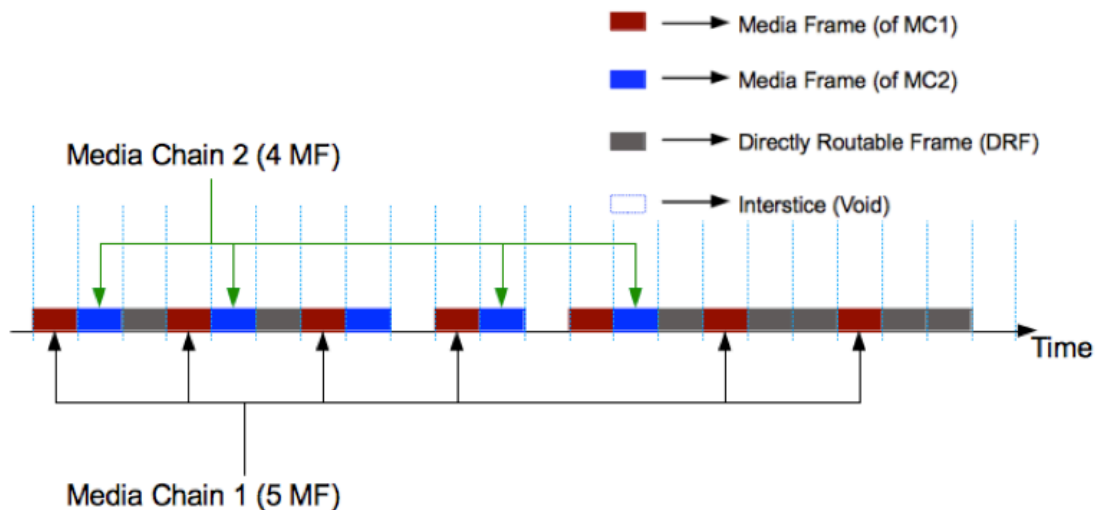


Figure 2.3 Two MCs, DRFs and Voids in a Channel [5]

On the other hand, the more numerous files with less than 10MB also occupy a significant fraction of bandwidth utilization. To facilitate the transferring of these files, the DRF (directly routed frame) is introduced. The DRF can be considered as an independent media frame, containing its own routing and management information,

which can be transferred through all the intermediate nodes directly. Since transparency may bring a large number of void interstitials between MFs, the DRFs can be placed into any of them.

2.3 Networking Methodology

To maintain the logical continuity of media chains and to minimize the buffering and switching costs, a transmission path must be scheduled in advance. Besides, it is also essential for the intermediate nodes to be able to identify the entire media chains, instead of inspecting each media frame.

2.3.1 Signaling and Control

Signaling is an essential procedure, used to establish and update network status, schedule an MC transfer, acknowledge receipt and many other functions. Two approaches are available for signaling: the first one is carrying the signal within media frame channel (e.g. in-band), by taking certain time slots between media frames; the second candidate is sending the signal in different channel, such as by the IP network (e.g. out-of-band). Considering the stability and flexibility of IP network, out-of-band signaling is more preferable at this stage. In this case, the intermediate routers (e.g. MAR, MBR) work upon both IP network and MF networks.

For manageability of entire network, a centralized server is assumed to be aware of and control the status of all routers. By retrieving the status information regularly, the central server can provide each router with global path, timing and occupation information. Moreover, the central server is also responsible for calculating the channel occupation and responding the channel reservation request sent from each router.

The peripheral devices, MAR and MBR are able to communicate with each other to exchange request information, and pass approval for requests from the central server. The requests are not guaranteed to be approved but the probability is high.

2.3.2 Media Chain Scheduling

Media chains are assembled and stored at the end user machines before transmitting. Once the MC is ready to transmit, a control packet (CP) is sent to the MAR first to schedule a channel. Typically, a CP includes at least the transparency degree (TD) of the MC, length of MC and destination/source address. MAR will respond quickly after negotiating with related intermediate nodes to establish and end-to-end path. An example of the negotiation is as follow: [5]

- 1) MAR obtains approval and path information from the central server, such as routing path and propagation time of each hop in the path.

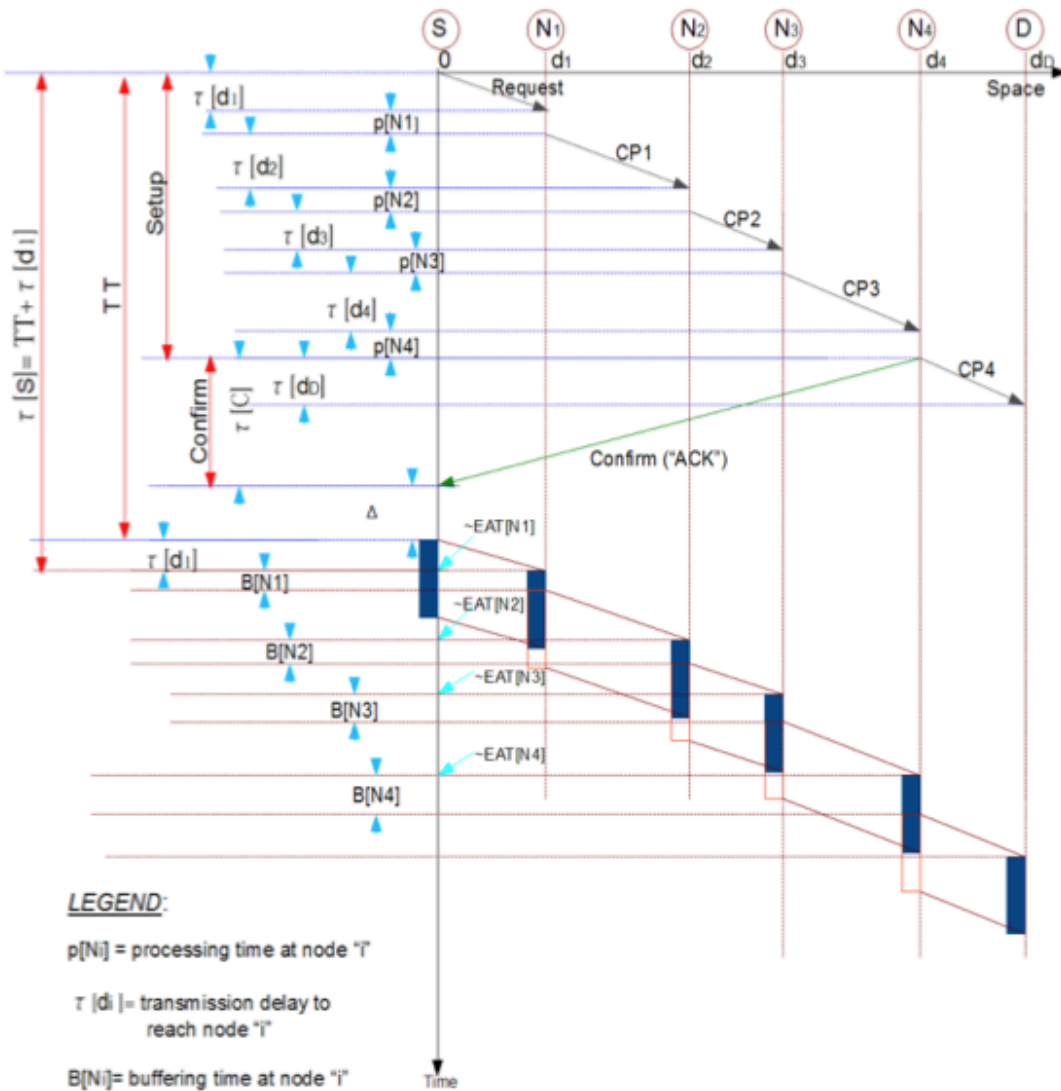


Figure 2.4 Path Reservation and Timing

- 2) MAR estimates a *Time to Transmit parameter (TT)*. This parameter is used to search for available time slot along the path by all the routers to schedule a MC transmission. TT is determined based on transmission delay from the user to MAR, hop distance along the path and processing time for the control packet at each node (Figure 2.4). The maximum number of requests that MAR needs to buffer is always less than the transparency degree, consequently the processing delay at each MAR is bounded by *transparency degree \times processing time* for one CP. In this figure, the quantity $\tau[s]$ is the time at which the ingress MAR should commit suitable resources. The quantify Δ takes into account the time uncertainties for the network hardware and $\tau[C]$ is the time for the confirmation to reach the source nodes.
- 3) MAR forwards CP to the next node along the path. The CP contains source/destination addresses, MC length, MC transparency degree, expected arrival

time ($EAT [N_i]$), and a unique ID related to MC/CP. The estimated arrival time indicates the time to wait after the reception of the CP at the nodes before the arrival of the MC.

- 4) The EAT is revised before the MAR sends the CP to the next node, due to the additional propagation and processing delay. This continues until the destination is reached.
- 5) After the reservation is successfully scheduled, a confirmation signal will be sent via IP network directly to the source node. Upon receiving the confirmation, the source node starts to transmit the MC at once. If the reservation fails, a “NACK” packet is sent though the reverse path to release the resources and the source node will resend the CP after waiting a random back-off time.

2.3.3 Synchronization

In MFN, channel reservation is scheduled by estimated arrival time, which is a hop delay along the path. Under this circumstance, a network-wide synchronized clock will be unnecessary. Local clock at each node is responsible for tracking the time difference between the reception of CP and the expected arrival time for the MC. Timing uncertainties can be compensated by buffers at each node, and they align the incoming MCs within the interstices of already scheduled MCs.

By eliminating the network-wide synchronized clock, the network operation could be simplified significantly. Moreover, removing the clock synchronization information among the network nodes saves some network bandwidth.

2.3.4 Directly Routable Frames (DRF)

In the case of sending smaller files, which may occupy a worthwhile fraction of one or just a few media frames, they will be sent as DRF. DRFs can be placed in any unoccupied time slot, including the interstices of MC. DRFs may or may not have the higher priority to transfer first due to the small size and flexibility, but they will probably be pre-empted by scheduled MCs. They are formed at the MAR or end users machine and contain a header for routing at each node.

2.4 Benefits and Advantage of MFN

The media frame network is expected to bring some attractive benefits. Primary advantages include:

- 1) Router Power Consumption

Currently IP lookup forms a bottleneck in packet forwarding due to the lower lookup speed compared to the high transmission rate in the IP routers, which also causes the linearly power consumption growth when the number of entries increases [20]. Some studies suggest that most power is consumed either in the switching function or header look up [21]. Lookup-consumed power dominates when packets are small. Another study provides sufficient evidence that by increasing the packet size, power consumption of IP lookup will be reduced significantly, because of the number of processed packets decreases. In the media frame network, the basic data unit is 10MB. Moreover, the media frames are concatenated as media chain in the transmission traffic, and only one header is needed for each MC. Consequently, the lookup rate in a media frame network will drop tremendously, which would slash the lookup power consumption to a negligible level. Study [22] indicates in large scale networks, the power consumption of header related functions is far larger than switching functions. Therefore, the media frame based network should bring large benefit on power consumption.

2) Efficient Channel Throughput

The media frame network is built based on very large container to transfer the files. The natural advantage is highly efficient throughput. Analogies in mail service, rail transportation and container shipping are convenient representations. Although the end-to-end reservation is not efficient for packet switching in an IP network, given that the MC consists of a large number of MFs, it is worthwhile to make a reservation. The periodicity of the MFs in a MC simplifies scheduling since an unoccupied time slot in one period is very likely to be available in the next period. In addition, placing a DRF in an available time slot brings further enhancement to the throughput.

3) Tight Coupling with Optical Transport System

In conventional IP networks, it is essential to maintain a dynamic capability in switching each relative small packet. But in the MF network, the input to output path is fully reserved, predictable and sufficiently stable over time scales substantially greater than with traditional router traffic.

4) Conformance with Existing Network Physical Layer

The MFN can work upon the current optical fiber network, by using a different switching and routing protocol. This advantage would reduce the costs largely when the MFN is widely distributed. The access network also share the same physical layer with existing standards, which follows in the next chapter.

2.6 Conclusion

In this chapter, the fundamentals of MFN are investigated. By introducing the hierarchical network, the topology of MFN is defined. Concepts of Media Frame and Media Chain are also studied. The MFN transfers data based on MCs, with channel reservation along the network. Scheduling and control mechanisms are briefly reviewed. DRF is introduced as an effective approach to transmit the files that are smaller than a typical media frame.

Several benefits and advantage of the MFN are presented. Due to the nature of MCs and MFs, MFN enables lower power consumption and efficient channel throughput. Because the MFN can be built upon the existing network infrastructure, the idea of MFN is economic and friendly to IP networks.

Chapter 3

Fundamentals of Access Network Design for MFN

In this chapter, the requirements and practical challenges of an access network for MFN are investigated. Referring to the current state-of-art technologies in access networks, a few options for MFN are presented. In October 2010, standards for XGPON, ITU-T Recommendation G.987, were finally finished, which offer us a basis on industrial specification to design and build the access network for MFN.

3.1 Challenges and Choices

As mentioned in previous chapters, a major challenge of MFN is to support MC transmission upon current widely deployed access systems. Two basic alternatives were presented. Networking media aggregation routers (MAR), and implementation of PON technology with a specifically designed transport layer. As discussed, the second alternative is more desirable for high-speed transmission and being able to assemble MF/MC at the client directly.

Current PON standards, including A/BPON, EPON, and GPON are all designed for ATM cells, Ethernet frames or IP packets, which obviously cannot accommodate MFs. Moreover, the current channel scheduling mechanisms are incompatible with media chains, because MC transmission requires a fixed bandwidth during the time in which the MC is transferred. Therefore, a series of protocols that specially accommodate MFs upon the PON's physical layer will be the major challenge in the development of an MFN access network.

Given the high capacity for both downstream and upstream, the XGPON standard defined by ITU-T recommendation G.987 is a natural candidate on which to develop the MFN access network. Since it coexists with current GPON components, XGPON will not require re-deployment of the network equipment.

3.2 Physical Layer Design

The physical layer, also commonly referred as PMD (physical media dependent) layer, is a considerable aspect during development of media frame network. Although on the first glance it seems to be a low-level engineering issue, it actually has tremendous impact on the performance of the network, such as compatibility of XGPON with fiber plant, system power budget, and range extension alternatives.

Table 3-1 Candidate Wavelength Bands and Potential Problems

Wavelength Channel	Potential Deficiency
Channel A 1595-1615 nm	Passive components not specified
Channel B 1540-1560 nm	Conflict with other system
Channel C 1530-1540 nm	Costs problem
Channel D 1340-1360 nm	Hard to achieve low loss
Channel E 1260-1280 nm	Occupied by XGPON

To overlay the MFN access network upon the XGPON system, the transmission for MFs may require a dedicated wavelength channel, in parallel with the existing IP services. The available wavelength bands had been listed and evaluated (table 3-1). Five wavelength channels were considered. Channel A was not a good candidate due to fears that fibers and passive components were not sufficiently specified for these wavelengths. Channel B is occupied by the video overlay. Optical components operating on Channel C would be costly because the GPON ONU may not have sufficient isolation to reject this wavelength. Optical fibers may be difficult to make with low loss on Channel D. Channel E was selected for the XGPON upstream channel. Hence channel C (1530-1540 nm) or channel D (1340-1360 nm) would be selected as they offer the least technical challenge. Alternatively, selecting from the guard bands in figure 1.11 may be an option.

For the splitting ratio, many commercial vendors build their ODN infrastructure with 1:32 or 1:64 split for their GPON system, hence 1:64 split shall be the minimum requirement for our MF-PON. Referring the generic splitter deployment in XGPON system in [15] (Figure 3.1), the splitting ratio can be extended to much higher levels. However, considering the maturity and cost-effectiveness of the optical devices, the selection on higher splitting ratio must be done with care.

Study [15] introduces the maximum fiber distance of at least 20 km, and the transmission container (TC) layer of XGPON needs to support the maximum fiber distance of 60 km. At this early stage of design of the access network, the reach distance supported by MF PON should be a minimum of 20 km.

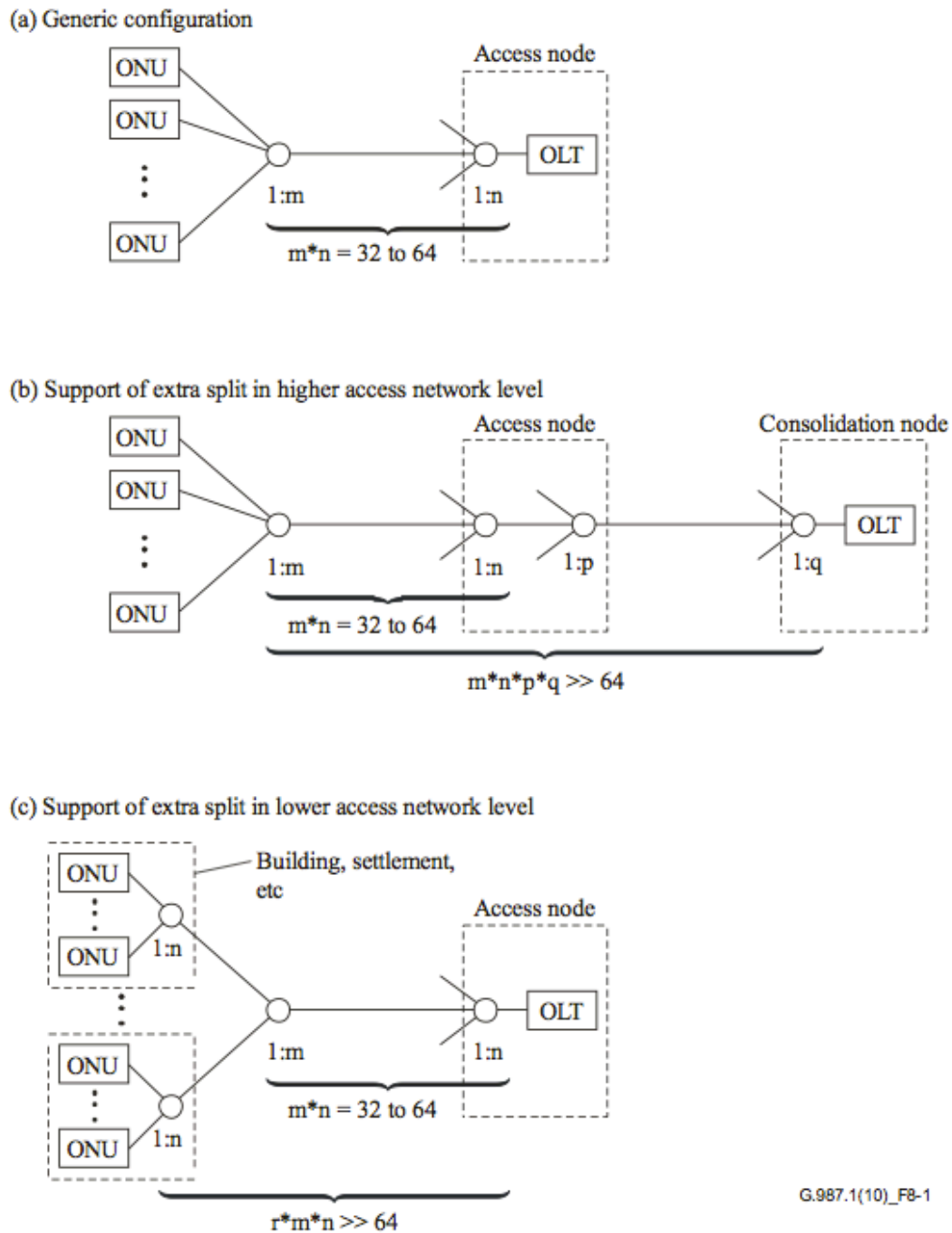


Figure 3.1 MF-PON Splitter Architecture Options [15]

Therefore, an MF overlay upon the XGPON system can be constructed by adding an MF-OLT before the WDM1r in the system, and the MF-ONUs are distributed at the users side (Figure 3.2).

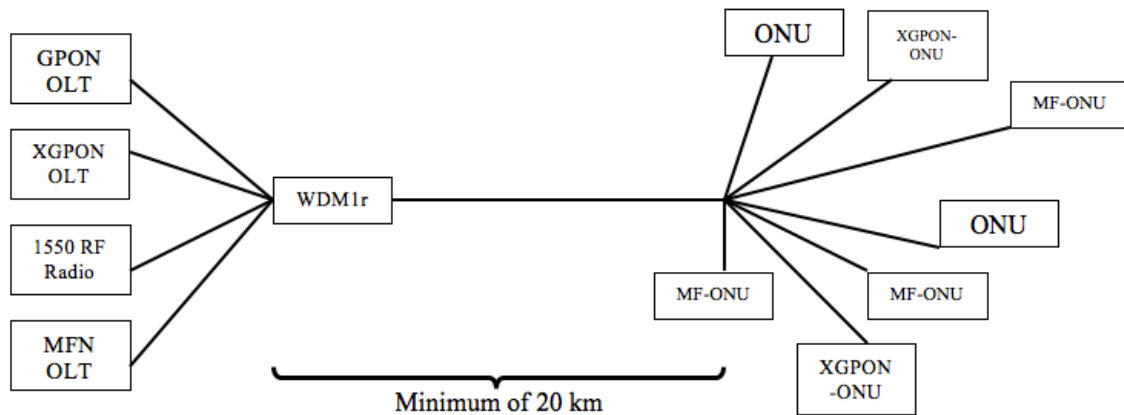


Figure 3.2 MFN Access Network Overlay upon XGPON System

In addition, there is also an option to enhance the system capacity by implementing WDM technology on the MF-PON system. By subdividing all the MF-ONU into different groups, each group can work on a separated wavelength channel (figure 3.3). In this manner, the supported number of end users can be tripled or more. Since each wavelength works independently, our protocol design is not dependent on the WDM overlay on the system. While aggressive for early stage of demand, this option does provide a convenient upgrade path.

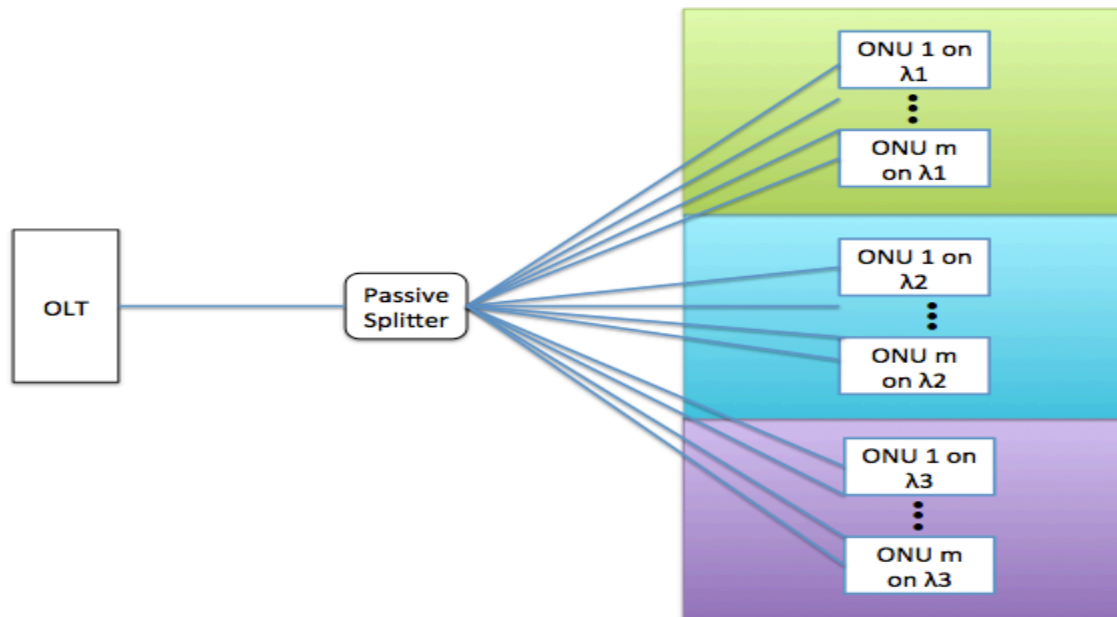


Figure 3.3 WDM on MF-PON

3.3 Logical Structure Design

Given a suitable plan for the physical layer, it was fairly clear that the MFN access network would work well on the basis of XGPON. However, since the data structure on MFN is very different with the conventional network, the logical structure of OLT and ONU, as well as the protocols, must be redesigned to accommodate MFs and MCs transport.

3.3.1 ONU Design

In the preferable scenario, the media frames and media chains are assembled at the end-user. In the case of FTTH deployment, each ONU is an end-user machine. Hence, the ONU is responsible for assembling and concatenating media frames, identifying the downstream traffic, and negotiating and reserving the path with the OLT for upstream.

The GPON ONU structure is as illustrated in figure 3.4. To support the media frame service, one option is to add a functional box between the modified PMD and Service MUX/DMUX, working as an overlay to realize all of the functions that are mentioned above.

As required for the reservation mechanism, a request for a channel grant should be sent to OLT first from the ONU. The control packet provides or conveys the request, and it contains necessary information such as source/destination address, media chain length and transparency degree. Note that there is no way to predict when the grant will be issued. Hence the control packet does not contain the estimated arrival time. This ETA value is a parameter used to reserve the end-to-end channel in backbone network. When the OLT negotiates with the intermediate nodes in core network, the granted starting time for media chain transmission from the ONU has been computed using ETA.

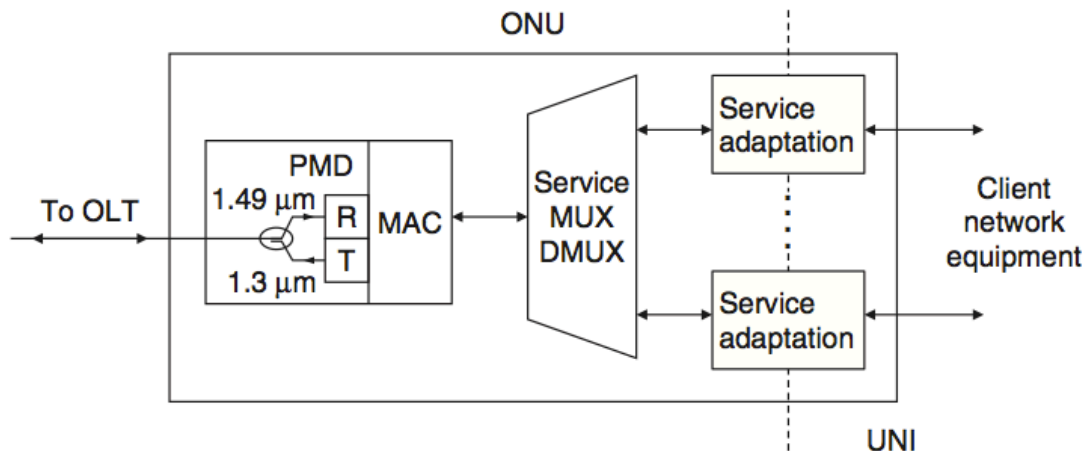


Figure 3.4 Original GPON ONU Structure

3.3.2 OLT Design

The OLT in a MFN access network serves in the central office, controlling and analyzing traffic conditions in the PON network, and making decisions based on ONU requests as to when to start transmitting the MF or MC.

Unlike the dynamic bandwidth allocation for IP packets, the media frame transmission of our PON is scheduled. The transmission speed standard for upstream is 10Gbps, and each time slot corresponds one media frame, which means the duration of each time slot is typically 8ms. The OLT should have full awareness of the slots of occupation of every 8ms-time slot. The OLT is responsible for processing the request from the ONU and passing the MC or MF traffic to the core network. To parallel these two tasks, the OLT should have separate input ports for requests and data. In the following chapter, a method for processing the requests will be presented. It requires a temporary storage queue before the request input port. The OLT also handles incoming traffic from the core network, which is shifted from, for example, higher transmission speed of 40 Gbps to a lower rate of 10Gbps. To accomplish this rate shifting, a buffer would be needed on the backbone side of the OLT to buffer the high-speed traffic. The size of this buffer is reduced dramatically as the transparency degree is varied, in this case by a factor of 4, between backbone and access. Figure 3.5 gives a general structure of the OLT.

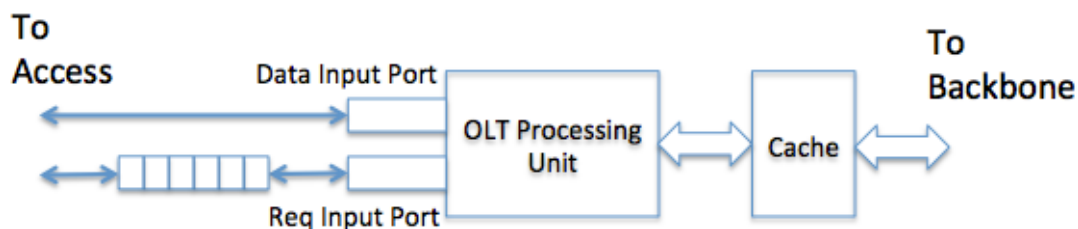


Figure 3.5 OLT Functional Blocks

In addition, as an important feature and advantage of the MFN, network-wide clock is not necessary. All the components in the access network only require a local clock to schedule the channel.

3.4 Conclusion

In this chapter, the physical layer and logical structures of the MFN access network are investigated. To achieve high capacity access, the last mile solution should be built based on the XG-PON system. However, all of the current existing PON standards are designed for IP network or ATM cells. Therefore, to facilitate media frames and media chains transport, the logical structure of network component and protocols must be reconsidered.

A wavelength plan was studied and selected for the upstream and downstream links. To achieve better performance, larger splitting ratios are preferred, and 64 should be the

minimum number. In addition, an aggressive WDM solution was proposed for enhancing the maximum user bandwidth.

In the functional structures of the ONU, a MF assembling block must be implemented to support assembling and disassembling the MF at the user side. At the OLT, a pipeline for the media chain requests should be built using the IP channel to transmit the requests that are required to schedule the MFs and MCs.

Chapter 4

Protocol Design

As mentioned in previous chapters, media frames cannot work on the IP network protocol due to its nature of path reservation and channel scheduling. In this chapter, the protocols that drive the MF PON system will be discussed and analyzed.

4.1 Initializing Procedures

Like other PON systems, an initialization operation must be conducted before the system starts working. In conventional PON systems, a synchronized clock must be employed for global scheduling to facilitate dynamic bandwidth allocation. But in the media frame network, this mechanism could be eliminated. Timing is based on the time difference between the reception of the grant signal and the start of transmission, and this time difference was referred as *SST* (scheduled starting time) in the previous chapter.

The initialization process in the access network establishes a database, storing the transmission delay from the OLT to each ONU, based on which the OLT can easily analyze the correct *SST* by subtracting the respective transmission delay. The data should be updated regularly due to the unpredictable characteristic of the network. Besides the ranging protocol, the following steps should be conducted in the initialization process:

- 1) The OLT broadcasts a time-calibrating packet to all ONUs, which contains the transmission start time at the OLT.
- 2) After receiving this packet, each ONU replies with a responding packet to the OLT, applying a random delay to avoid conflict in the shared fiber. These responding packets contain the ONU ID and the random delay time.
- 3) The OLT analyzes each received responding packet by subtracting the delay time and restoring the result as *RTD* (round trip delay). This *RTD* will be used to generate the channel grants for the ONUs and give the *SST*.

4.2 Channel Sharing Schemes

In the PON system, the fiber channel is shared between all ONUs on the same wavelength. A bandwidth allocation protocol must be established to guarantee that the system works in order. Because of the logical continuity of each media chain, the channel occupation cannot be broken after the scheduled starting time. Since the transparency mechanism has been included, several media chains could be transmitted simultaneously. Therefore, the configuration of transparency has a very significant effect on the system performance. A few configuration options will be discussed in what follows.

4.2.1 Fixed Transparency Degrees

In this scenario, every media chain sent from/to the ONU has the same transparency degree. Technically, the transparency degree value could be any integer. However, in order to be compatible with the doubled or halved transparency degree (this will be discussed in the later phases), it is always better to limit the transparency degree to a power of two. Under the PON physical standards (10Gbps down/up), the effective line speed of transmission for a given user can reach 10Gbps only when the transparency degree is one. When the TD is doubled, the effective line speed will be halved. To maintain an acceptable transmission rate, the TD will be kept no larger than 16 in the design stage. Figure 4.1 simply illustrates this fixed transparency degree of 4.

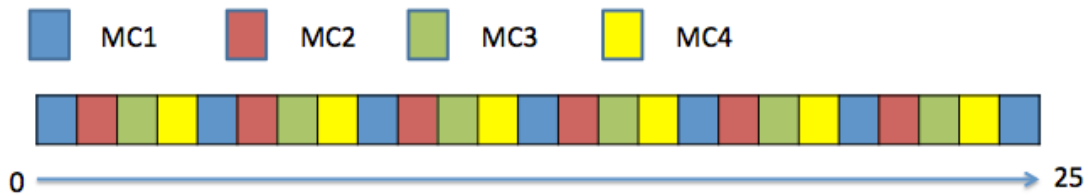


Figure 4.1 Illustration of Fixed Transparency Degree of four

The advantage of a fixed transparency degree is obvious: all the media chains share the same fraction of the channel, and effective line rates are the same. On the other hand, the maximum supported users at the same time are limited to the number of transparency degree. If this TD number increases, the effective line rate correspondingly drops. This trade-off must be considered to reach a balance between maximum number of supported users and line rates.

4.2.2 Variable Transparency Degrees

Note that the variable TD does not mean the transparency degree varies in one media chain, but different MCs could have a different TD number. In the research stage, we define the maximum TD as 16, and the minimum as 2. The TD is varied among 2, 4, 8 and 16 to accommodate different number of requesting users and channel utilization conditions. Figure 4.2 illustrates this scheme. Different MCs with different TD are interleaved at the time domain, and the white slots represent the unoccupied times. Comparing to the fixed TD scheme, a not very obvious disadvantage emerges, which is that some time slots may be wasted. From the figure 4.2, a few available white slots can only be reserved by some certain TD, or they will be wasted.

To allocate bandwidth in the case of variable transparency degree, a method must be established to guide the allocation process. In this thesis, a reference parameter is introduced, *CAD* (channel availability degree), to mark the occupation condition of the channel. Here, the maximum TD is 16, so we define the maximum *CAD* as 16. The time channel can be seen as a sequence of 16-slot periods in the time domain. When a TD of 8 is granted, two out of the 16 time slots will be occupied, hence we subtract two from the current *CAD*. Similarly, we subtract 8 from current TD when TD of 2 is granted, and

subtract 4 when TD of 4 is granted, and subtract 1 when TD of 16 is granted. Under this case, the system can support up to 2 MCs with TD of 2, 4 MCs with TD of 4, 8 MCs with TD of 8, and 16 MCs with TD of 16. An optimized combination of these different TD users should be selected (Figure 4.3) to achieve a balance between the maximum supported users and line rate for each user. We can assume the ONUs have different priorities to generate different TDs based on their class.

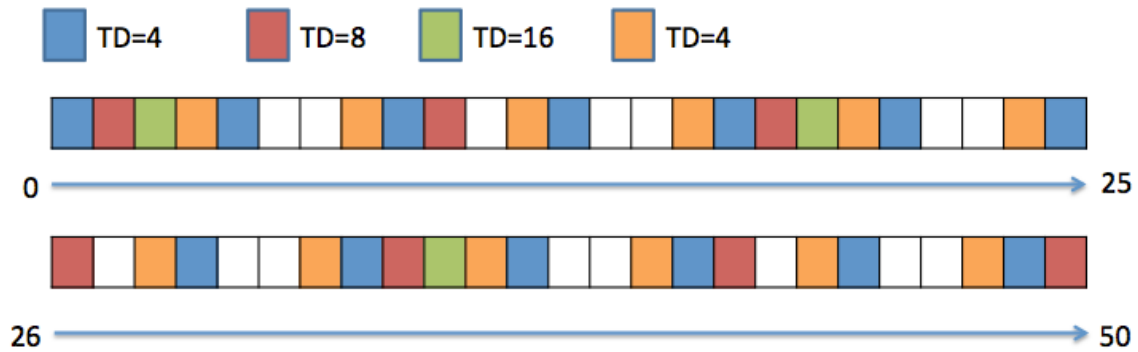


Figure 4.2 Variable Transparency Degrees

Figure 4.3 shows a combination of the variable TDs. If there is one active transmission with TD of 2, only 4, 8 and 16 can be granted. If there are already three active transmissions with TD of 4, only 8 and 16 can be granted. When six transmissions with TD of 8 in the channel, only 16 can be granted. A trade-off must be made between the number of users supported and the transmission rate for each media chain.

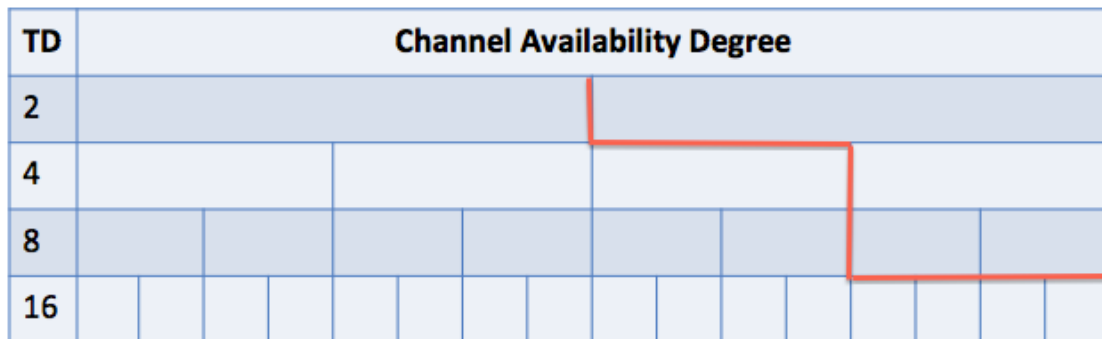


Figure 4.3 One combination of maximum numbers of TD

4.3 Negotiation Protocols

Before the transmission of media chains starts, the ONU negotiates with the OLT to schedule a transmission channel in the PON system. In this design, the control packet serves as the request for reserving the channel. The OLT analyzes the information of current and future channel occupation after reviewing the requests from the ONUs. If

certain time slots are available and suitable for the requesting MC, the OLT will grant the channel reservation, otherwise OLT will reject the request for this time.

The media chains for uploading are assembled at the ONU, and the request is sent to the OLT. For downstream transmission, the OLT receives media chains from the core network, and then inserts this into the downstream traffic. Both ONU and OLT need a storage cache for temporarily storing the media chains. Two different ways of negotiation for upstream traffic are presented in below, which mainly differ in how to further process when the request is rejected.

4.3.1 Request Resending Based Negotiation

This method is illustrated in the flowchart in figure 4.4, and can be described as following steps:

- 1) ONU sends a request to OLT, which includes the requested TD, MC length, and priority information (if necessary).
- 2) Upon the arrival of the request, OLT examines the *CAD* value, to check if *CAD* is enough for granting the requested TD. If *CAD* is not enough, the process jumps to step 3. If *CAD* is enough and there is no limitation on the quota for the specific TD (referred in Dynamic Transparency Degree paragraphs), it proceeds to step 4.
- 3) OLT replies to the ONU with a “resending” signal. After receiving this signal, the ONU resends the request after a random delay. Then step 1 repeats.
- 4) OLT replies to the ONU with a grant signal, containing the scheduled start time and granted TD value. Meanwhile, OLT subtract a corresponding value on *CAD*, to mark the channel occupation condition.
- 5) ONU starts the MC transmission after the scheduled time difference. After receiving the entire MC, OLT adds back the subtracted *CAD* value.

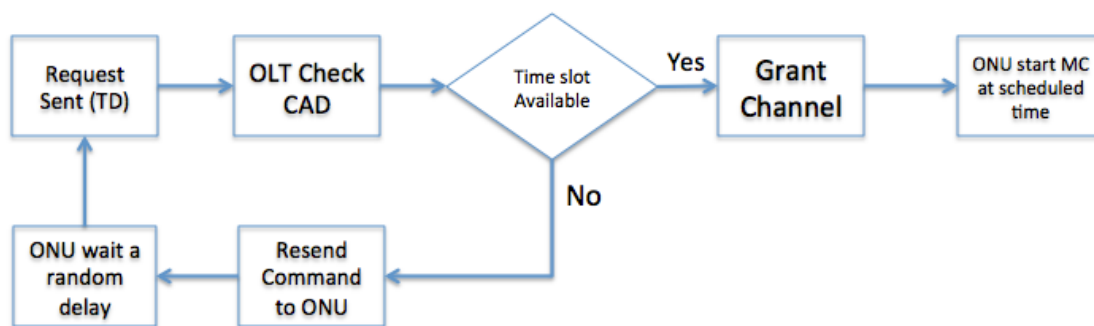


Figure 4.4 Request resending based negotiation flow chart

For DRFs, the OLT will not reject requests. By searching along the time axis, the DRFs will be scheduled in the earliest available time slots to assure the flexibility and priority of DRFs.

This method is advantageous in that it provides fast processing and easy to implement. The random delay after being rejected can reduce the rejection possibility for the next time. However, it is still possible that some “unlucky” ONU may step into a requesting loop, causing a long waiting time. In addition, the repeating system requests reduce the efficiency especially under the higher traffic conditions. To remedy this defect, another method called queuing based request processing is presented below.

4.3.2 Queue Based Request Processing

In the queue based request method, a FIFO (first in first out) queue is placed before the OLT. All requests will get into the queue, waiting the OLT to grant the channel. If the queue is empty, the request will be directly delivered to the OLT, otherwise it will wait in line.

Figure 4.5 demonstrates the steps:

- 1) ONU sends a request, which enters the queue from the tail.
- 2) If the queue is empty, it serves this request to OLT. If the queue is not empty, it waits for a signal that triggers the queue serve the first request to OLT. In this design, only a MC finishing transmission signal can trigger the queue serve the first entry.
- 3) OLT analyzes the CAD, checking if it is enough for granting the requested TD. If not, this process leads to step 4. If yes and there is no limitation on the quota of the TD, it proceeds to step 5
- 4) OLT puts the rejected request back in the first place of the queue. It keeps waiting in the queue until a MC finishing signal triggers the queue. Then the request in the front of the queue gets back to the OLT and repeats the step 3.
- 5) OLT replies the ONU with a grant signal, containing the scheduled start time and granted TD value. Meanwhile, OLT subtracts a corresponding value on CAD to mark the channel occupation condition.
- 6) ONU starts the MC transmission after the scheduled time difference. After receiving the entire MC, OLT adds back the subtracted CAD value. In the meantime, OLT controls the queue to serve the first request.

When a DRF arrives, it will be directly delivered to the OLT. The OLT will search along the time axis for a time slot to fit this DRF. This could be also implemented to handle the priority classes of MC requests.

By placing the upcoming requests in a queue, the number of system requests drops. In request resending based method, the total request number equals $valid\ request\ number / (1 - blocking\ probability)$. Whereas in the queue based method, each request will be sent only once, which enhances the channel utilization by eliminating the unnecessary requests. All requests will be processed in the order of arrival. In this case, the “unlucky” situation mentioned in request resending will not happen, since the rejected request will wait in the front of the queue until the CAD is available for the proposed TD.

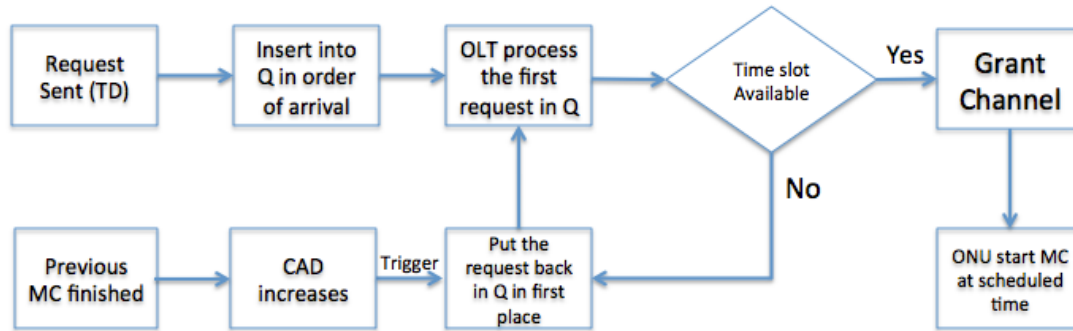


Figure 4.5 Queue based request processing flow chart

In summarizing these two request process methods, it can be found that the OLT always grants the TD that ONU requests. But in some cases granting a reduced TD may bring benefit both to the requesting ONU and to the other queuing requests. For example, if a request with TD of 4 comes to the OLT, and CAD can only support a maximum TD of 8, the options for the request are either to wait in the queue for CAD available for 4, or to change the TD to 8. However, decreasing the TD to 8 will cost twice the original transmission time, but it is worthwhile if the original transmission time plus waiting time is longer. Therefore, an analysis on time-cost comparison is desirable before the OLT makes decisions.

4.3.3 Analysis on Time-Cost Comparison

This comparison is performed by the OLT before making decisions on rejecting the requests or increasing TD, depending on which consumes less time from the moment the request is received to finishing transmission. Table 4.1 displays which TD case to compare in this analysis.

Table 4-1 Cases to Compare

		Requested TD			
		2	4	8	16
Candidate TD	2	No	No	No	No
	4	Yes	No	No	No
	8	Yes	Yes	No	No
	16	Yes	Yes	Yes	No

In the above table, each TD is compared only to those with larger TDs. TD of 16 will not go through this analysis since it is the largest TD in the design. There are many ways to compare the time cost, for example, as shown in figure 4.6.

- 1) Upon reception of the request, OLT checks the CAD. If it is available for granting the proposed TD, then OLT grants the channel, otherwise it proceeds to step 2.
- 2) According to the CAD and requested media chain length, OLT estimates the time cost of the media chain going by a grantable (best available) TD, recording that as t_1 .

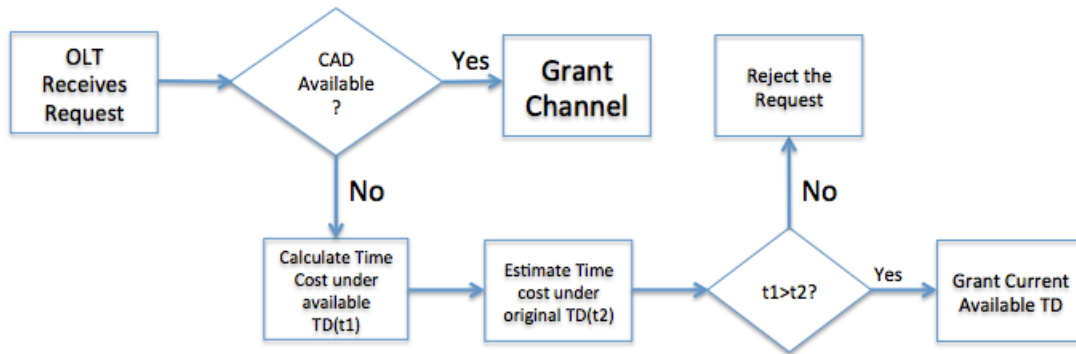


Figure 4.6 Flow chart of time comparing

- 3) According to the current channel reservation schedule, OLT searches along the time axis for available slots that fit the requested TD. Then OLT calculates the waiting time plus transmission time under the requested TD, recording that as t_2 .
- 4) By comparing t_1 and t_2 . If $t_1 > t_2$, the OLT will grant the ONU with current available TD instead of the requested one. Otherwise the OLT rejects it, and the request goes back to the front of the queue.

Note that even if the current *CAD* is zero, the available time slots could always be found if the OLT searches along the time axis. The reason that the OLT does not schedule future time slots is to avoid over-scheduling. If time slots in the far future are reserved, DRFs cannot be scheduled and transmitted in a short time. The number of DRFs cannot be neglected because DRFs may take a significant portion of all requests.

With this optimization, the OLT is able to make better decisions on the requests. The congestion problem before the OLT could be alleviated. Numerical analysis and simulation results will be presented in next chapter.

4.3.4 Downstream Traffic

The designed data rate in the core network is typical 40Gbps, which is four times of the access network. In this case, the transparency degree of the media chains in the core network should be four times that of the access network. Therefore, media chains with TD less than 4 from the backbone will not be used. Since the interface between the MF

access network and the MF core network has not been explored previously, we investigate some specific cases.

If the end-to-end reservation in the media frame network is actually end-to-end, which means the data channel is reserved from user machine to another user machine or server, downstream traffic congestion will not be of concern. All possible conflicts and congestion will be considered by the central state server before establishing the channel reservation. But if potential congestion at the access network is not considered by the central state server, the OLT must take the responsibility to resolve the conflicts.

4.4 Conclusion

In this chapter, the channel reservation protocol is investigated. By presenting the simplest protocol-request resending based process, the advantage and disadvantage of this method are discussed. The most critical defect is too much unnecessary repeating requests and rejecting signals. To solve this problem, a second option was proposed. By simply adding a queue before the OLT processing unit, all incoming requests are pushed into the queue. Each request is processed in the order of arrival. If the request is rejected, it will be sent back to the front of the queue, waiting for sufficient *CAD* to grant the request.

Aside from the processing protocol, an optimization algorithm is also presented to enable the OLT to make better decisions. When the current *CAD* is unable to grant the requested TD, the transmission time with the current maximum supported TD is compared with the waiting time plus transmission time under the original TD. The one costing less time will be selected and the OLT will give a decision.

In addition, downstream traffic is also discussed. Data traffic gets shifted by the TD to lower speed from higher speed to accommodate the lower rate of the access network.

Chapter 5

System Performance Analysis and Simulations

In this chapter, the system performance is estimated by some key examples. By implementing the reservation protocols and algorithms referred to in the previous chapter, a module-based simulation system was built to examine the performance on different traffic conditions. Given the more challenging requirements for collision avoidance in the upstream, only the upstream channel is considered.

5.1 System Performance Indicators

To achieve the best performance, the system should reach a balance between two factors: minimizing the time cost for transferring media files and maximizing the number of supported users. On this basis, the waiting time in the request queue should be shortened while the channel utilization should be maintained at a high level. Several values that will be measured in the simulation are listed.

Load (requests/sec): Given the difference between simulation and practical operation, it is impractical to specify the behavior of every single user behavior in the simulation to generate different traffic conditions. Fortunately, by controlling the frequency of all users sending large files, a series of equivalent traffic load values can be obtained at the shared fiber channel.

Round Time (Total Time): The most essential factor to rate the system performance from the users perspective is how long it will take from clicking the “send” button to finishing the transmission. This time includes the request waiting time, transmission time and channel delay. However, given that each media chain has different length, individual round time values may not be able to reflect the system performance. A normalized value could be a reference, which is dividing the total time by the media chain length.

Queuing Time: This value describes the request waiting time in the queue before being granted. This time cannot be too long because the streaming media transfer represents a significant fraction of the large file transmission, and a basic requirement of streaming media is starting transmission as soon as possible. If the queuing time were too long for a transmission, it would be inefficient for streaming media.

Queue Length: This value can be a direct indicator of the congestion situation.

Channel Utilization: This can be obtained from the ratio of the number of occupied time slots and the total number of time slots. This value indicates the channel efficiency on transporting the media frames.

5.2 Simulation System

A MF access network was constructed by the Omnet++ simulation package to emulate the MC reservation protocols under various traffic conditions. Each ONU represents a user generating upload traffic with an exponential distribution and variable mean as the inter-arrival rates of requests.

The simulation topology is shown in figure 5.1. Each ONU sends request as a control packet to the OLT when a media chain or DRF is generated. The optical power splitter/combiner passes the upstream traffic and duplicates downstream traffic to all ONU. To emulate the practical media chain traffic, it is presumed that the media chain length is uniformly distributed between 100 and 1300, which represents the typical media files with size between 1GB and 13GB. Note that the upstream capacity is 10Gbps, and the average MCs size is 7GB. Among the requests, it is assumed a portion of 20% is for DRFs, of which the size is negligible comparing to the MCs. Therefore, the upper bound of the system load can be calculated: $10\text{Gbps}/(7\text{GB}\cdot 80\%)=0.223$, where the 10Gbps is the channel capacity. So in all following simulation figures, the load value converges at the point of 0.223.

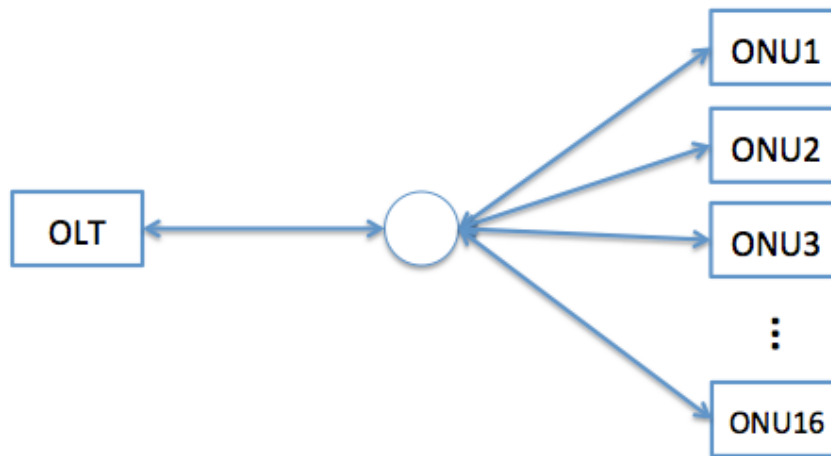


Figure 5.1 Access network topology for simulation

Omnet++ is a module-based network simulation tool. To run the simulation, three modules were developed: ONU, OLT and splitter. The ONU is responsible for sending requests and the OLT is in charge of granting the channel to ONUs. The splitter duplicates the message from the OLT to all downstream ports, and passes the message from ONUs to the upstream port. The interval time between two requests from one ONU

is exponentially distributed. By adjusting the mean of the exponential distribution, different traffic load can be generated. In the simulation, the mean is set in the range of 0 and 200 seconds, with step of 5 seconds. For each TD scenario, the simulation runs 41 times to obtain the performance results at different traffic load.

For each traffic load condition, we can obtain accurate performance results when the queue at the OLT maintained at a stable length. We configure the simulation to run for 5 minutes after a stable length has been achieved, and use the average values of all the transmissions as the performance results. Note that the time units here are simulation time rather than real-world time. In the simulation, the time is scaled by 8ms, which is much slower than the real word time.

5.3 Simulation for Fixed Transparency Degree

The simulation was started from the simplest scenario, fixed transparency degree of one, and then extended to larger fixed transparency degree. By comparing the indicators, we can draw conclusions as to the difference between TDs.

5.3.1 TD of 1 Media Chains

By setting the transparency degree to one, the transmission reaches the full speed of XGPON (e.g. 10Gbps). However, this creates difficulties on the DRF reservation as only one MC can be sent at any time.



Figure 5.2 Media chains with TD of one

From figure 5.2, it is obvious that in the MC transmission, there is no void slot during this time. Under this case the flexibility feature of media frame network is compromised.

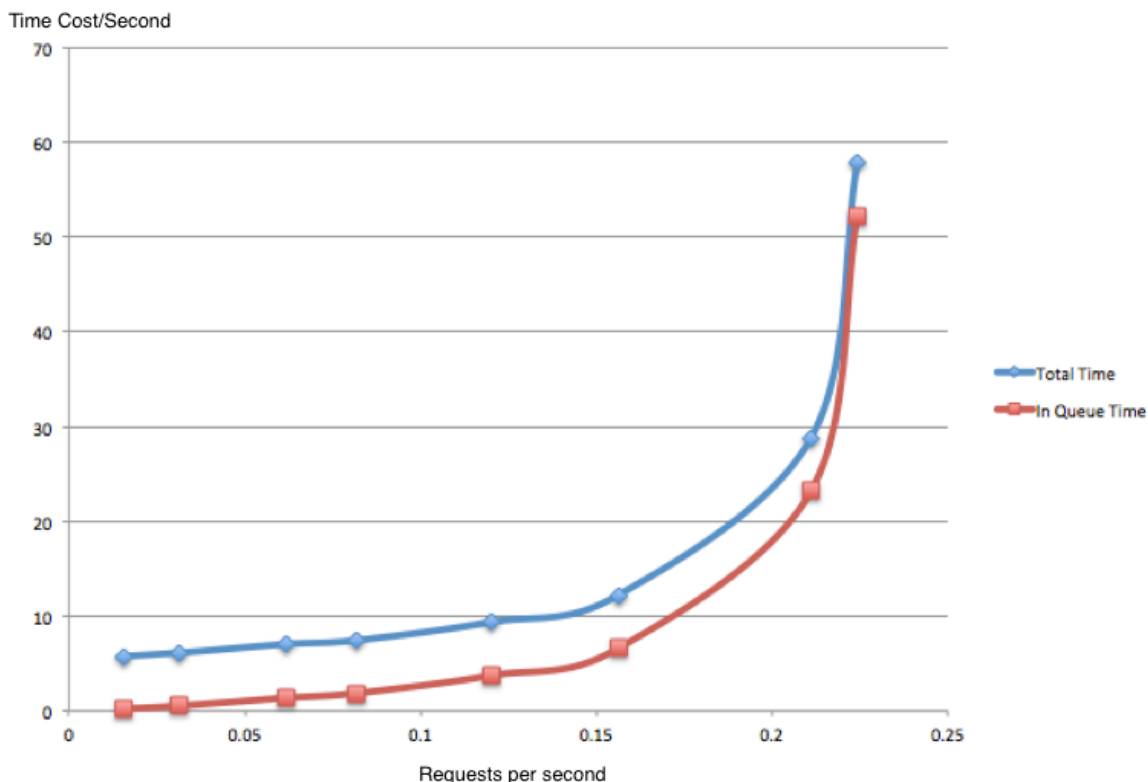


Figure 5.3 Total time and in-queue time for TD1

Figure 5.3 shows the results of transmission performance of this scenario. Apparently, the transmission speed is very fast especially under lower traffic condition. However, the problem appears for higher traffic conditions where most of the time is wasted on the requests waiting in the queue. At the load of 0.23, the round time average is 58 seconds, while the requests spent 52.3 seconds in the queue before being granted. This is decent for transferring very large files, but for the users of streaming video, the time distribution is not acceptable.

Figure 5.4 presents the channel utilization and blocking probability. We predefine a reference parameter here: a threshold of 1000 milliseconds for the waiting time. Requests that stay in the queue for a time exceeding this value are considered as false blocked (or delayed). In fact, with the waiting queue in the OLT, “blocking” will never happen, and the request can be only delayed. Therefore, a blocking probability obtained here is only seen as a reference. Note that one request may be blocked several times before it is approved by the OLT. With TD of one, the channel utilization and blocking probability increase linearly as the system load grows.

As designed in the provisions, the DRFs in this case have highest priority. They will be scheduled and sent as soon as available time slots appear.

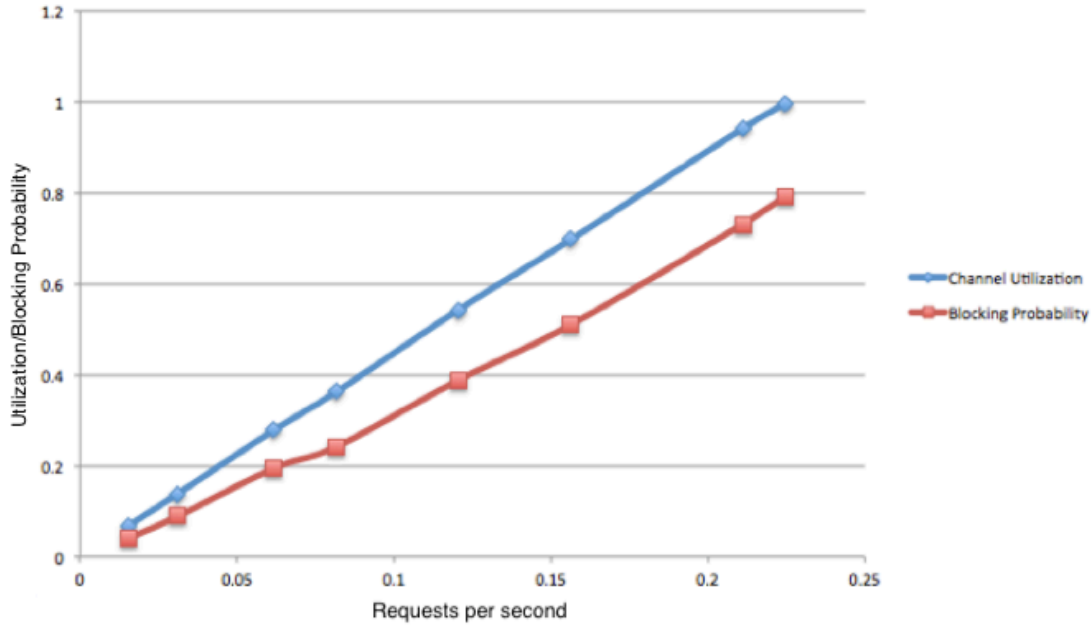


Figure 5.4 Channel utilization and blocking probability for TD1

5.3.2 Larger fixed TD Simulation

After discovering the advantage and shortcomings of TD1, we extended the simulation to large transparency degree numbers. Here only two cases are simulated, TD2 and TD4. For higher TDs, the results show simulation patterns, but with lower traffic round time.

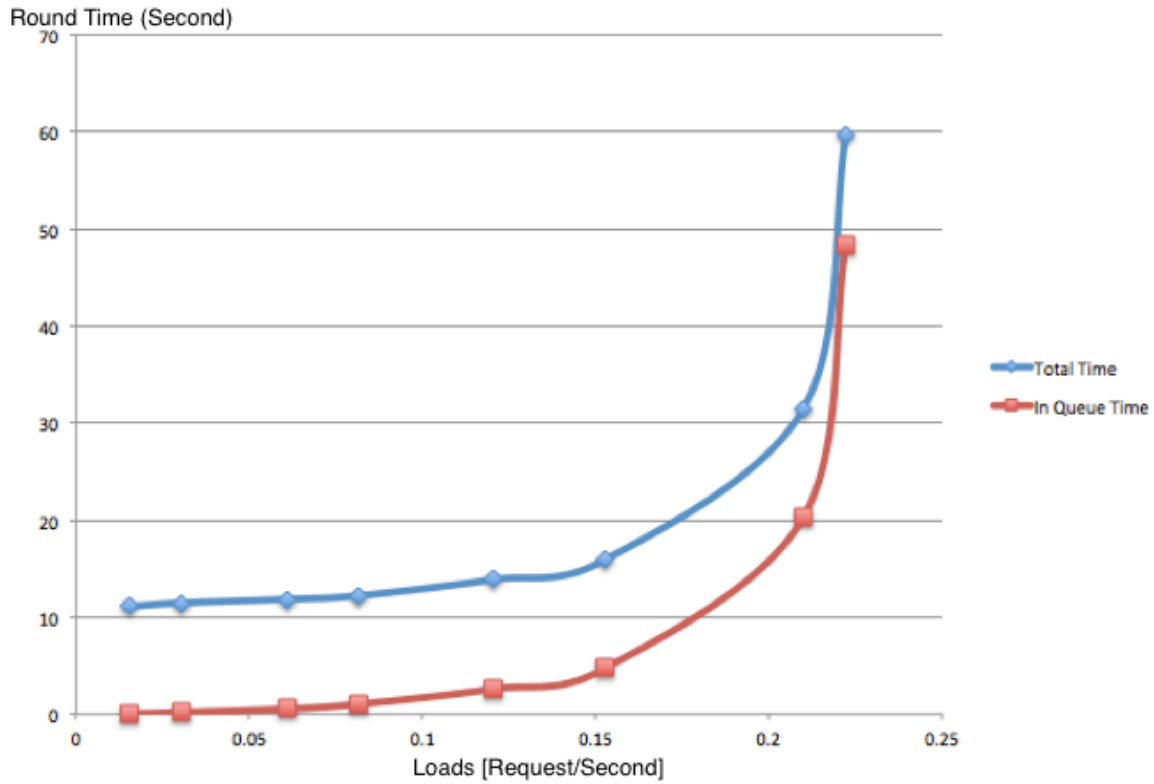


Figure 5.5 Total time and in-queue time for TD2

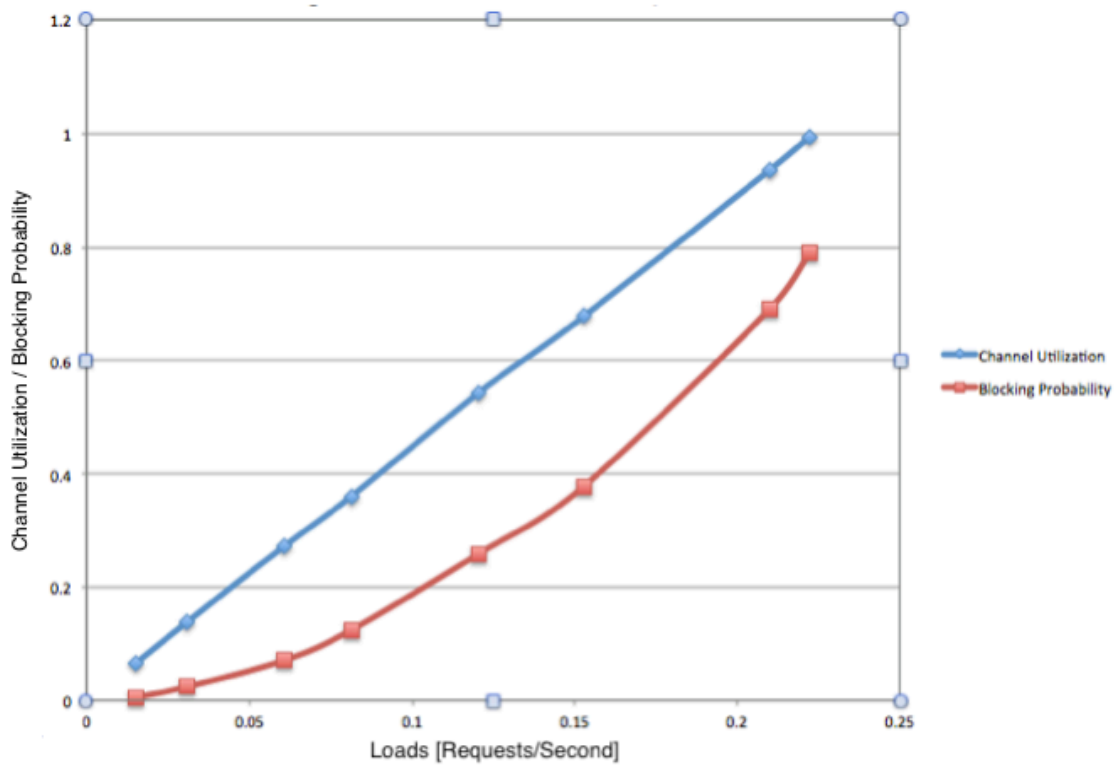


Figure 5.6 Channel utilization and blocking probability for TD2

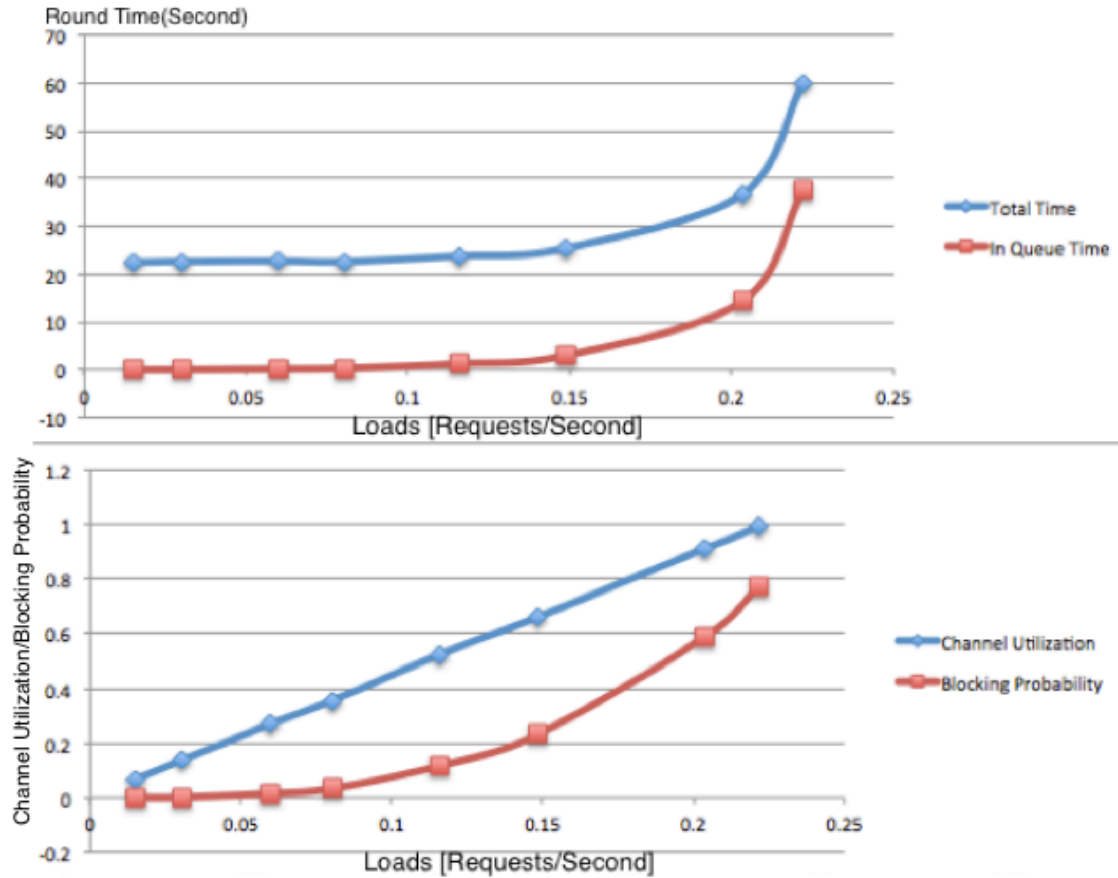


Figure 5.7 Results for fixed TD4

Figures 5.5-5.7 display the system performance with fixed TD2 and TD4. From these figures we can conclude that systems with fixed TD have similar patterns on performance. As the TD rises, the total time costs almost the same at higher traffic condition, but it costs more at lower traffic. A comparison of the average queue length comparison at higher-level traffic is presented in figure 5.8.

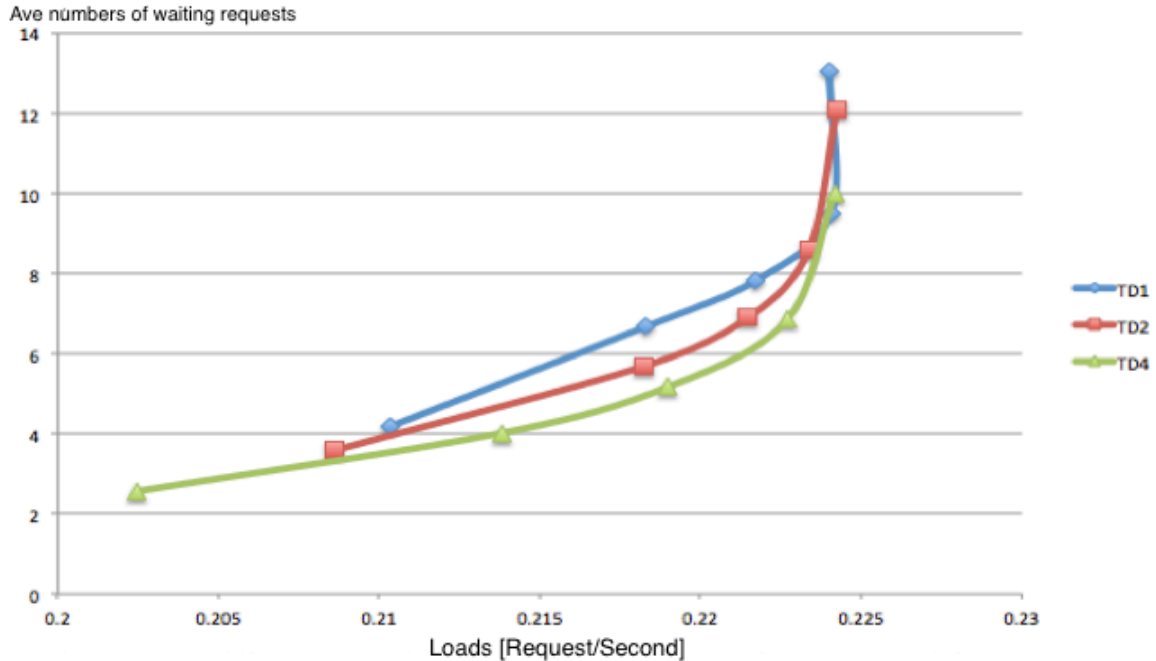


Figure 5.8 Queue length comparison between fixed TDs

As a further extension, the round time (total time) and waiting in queue time of fixed TD of eight (figure 5.9) are given.

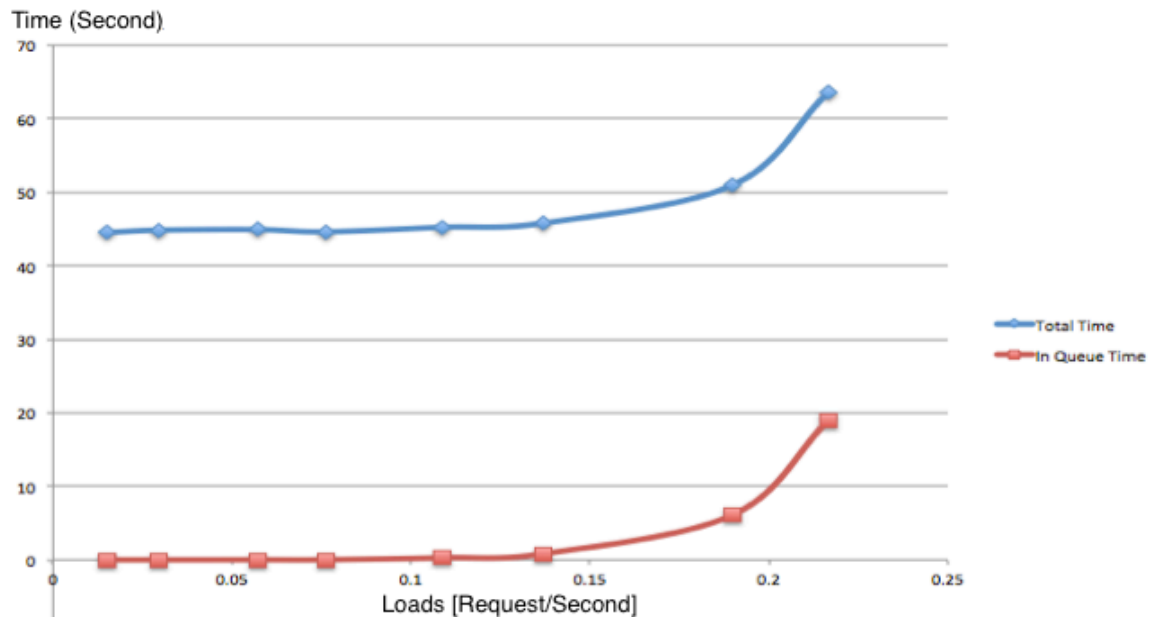


Figure 5.9 Time values for fixed TD of eight

By comparing the round time in figures 5.3-5.9, it is fairly obvious that at lower traffic, smaller TD can finish the transmission faster because the number of requests is small, and each media chain has the chance to transport at full speed. Under higher traffic,

larger TD requests would wait less time in the queue, while the total time costs are almost the same among different TDs. Therefore, it would be a feasible solution to implement different TD at different traffic conditions, in order to optimize the time cost on sending files.

5.4 Analysis for Variable TDs

Technically, enabling variable TD could bring some flexibility to the system bandwidth allocation, and enhance the channel utilization. For instance, if a request with a TD of four is waiting in the queue, the CAD only supports TD of 8. It is very likely that the request will wait for a long time to get the reservation. However, on the other hand, if the request is changed to the TD of 8, it can be granted immediately. The round time cost at second condition may be less than that of the first.

5.4.1 Combination of TD1 and TD2

As the first step, only TD of one and two are selected as allowed transparency degrees in the system. The ONUs generate requests with randomly selected TD. In the OLT processing, TD of 1 can be changed to two, based on the round time cost analysis.

It is expected that the performance of this approach will be worse than all fixed TD scenarios at high level traffic, by comparing the transmission time, waiting time and channel utilization, but it would be much better at lower traffic. Results showed in figure 5.10 correspond this deduction. When the channel is busy, the best solution to eliminate the request congestion is sending out the media chains and releasing the channel as soon as possible, and other optimizations only have limited effects. Comparing with simulation results of TD of 1 and other larger TDs proves this deduction.

5.4.2 Max TD of four

In this step, we add transparency degree of 4 into the allowed TD range. More users will be supported at the same time, so the number of waiting requests in the queue can be reduced. In this case, we run the simulation on two different selections of TD. The first group is TD 1, 2 and 4, and the second selection is only 2 and 4. Results are in figure 5.11.

From the graphs, it can be seen that the TD of one enhances the system performance at low traffic, but brings serious delay at high traffic. Meanwhile, the channel utilization maintains approximately the same level at each traffic condition, which is illustrated in figure 5.12.

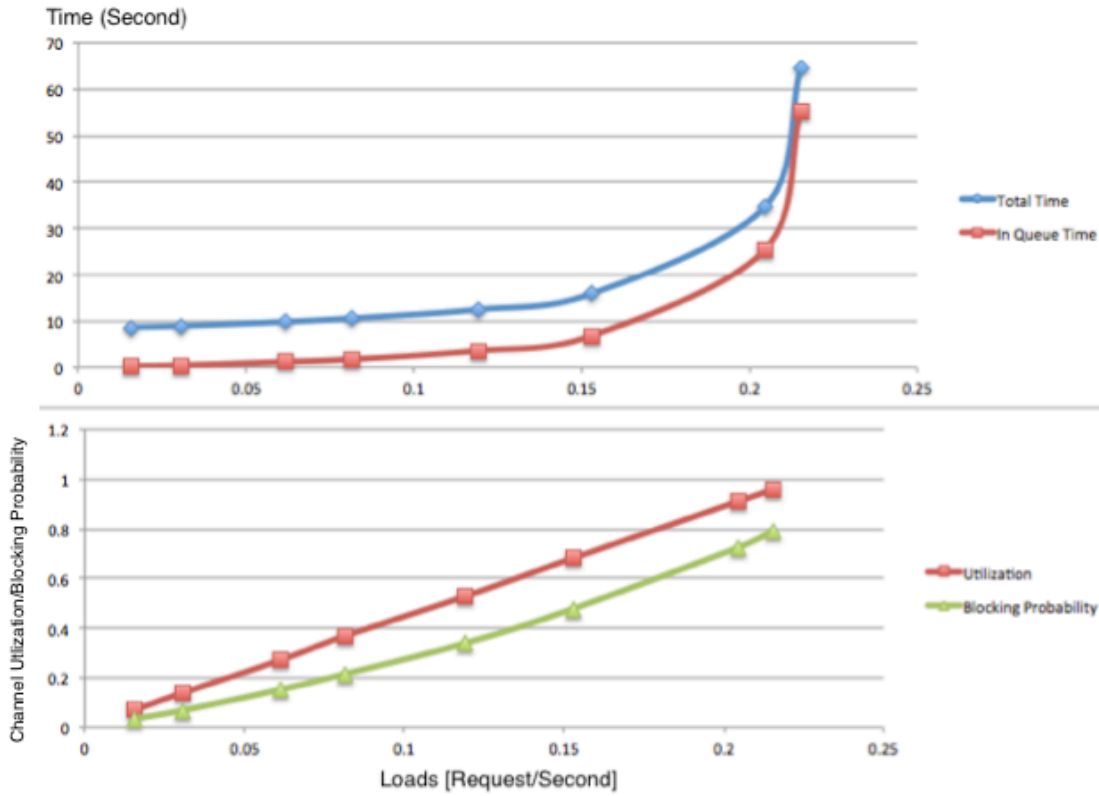


Figure 5.10 System performance of TD1 and TD2 combination

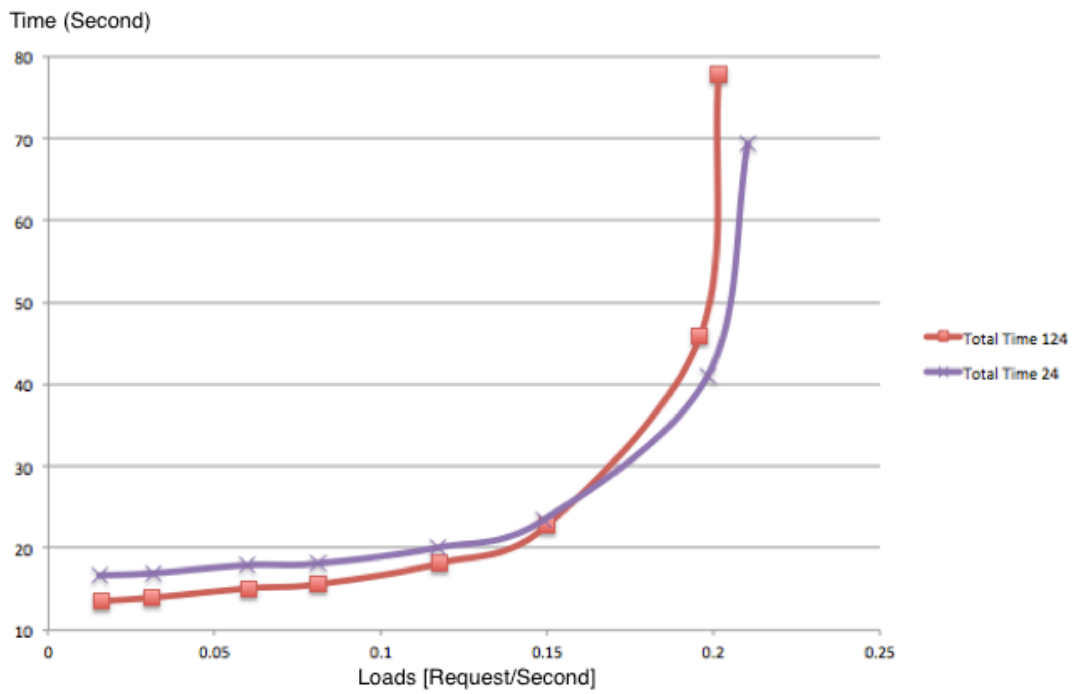


Figure 5.11 Total time comparison between TD124 and TD24

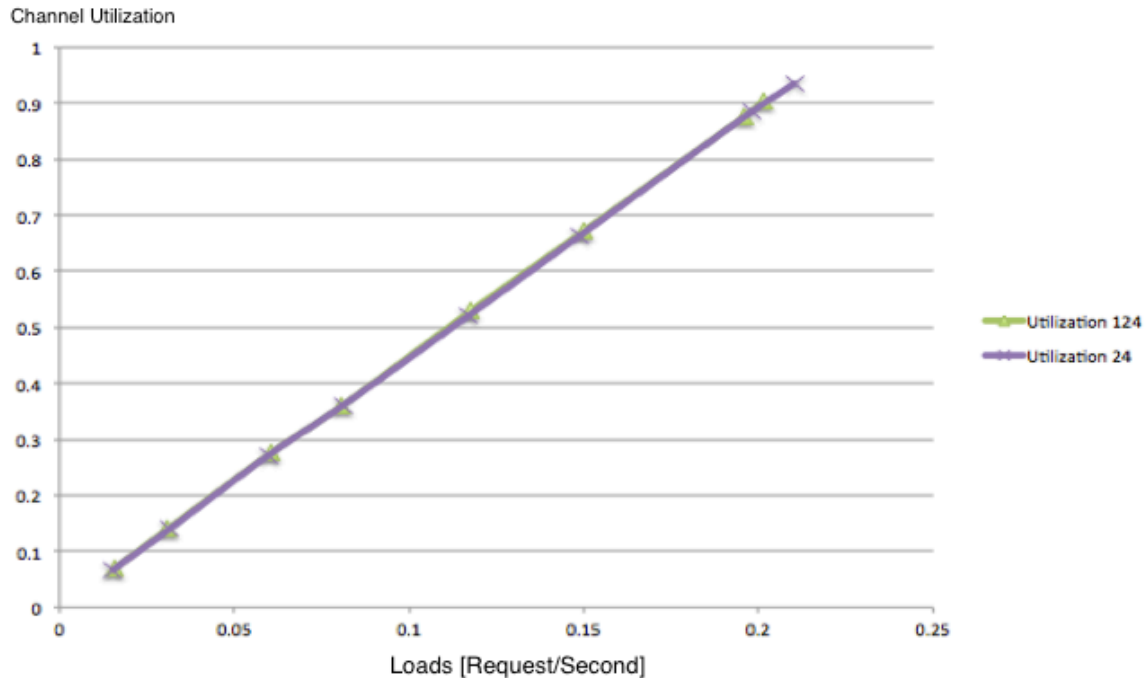


Figure 5.12 Utilization comparison between TD124 and TD24

5.5 Summary

By summarizing the numerical results in this chapter, a conclusion is reached: larger TD accommodates the high traffic in the system, while smaller TD brings less round-time cost. Using a fixed TD of one brings efficient transmission at a high rate, but constrains the transmission or insertion of DRFs. Larger fixed TDs are more friendly with DRF reservation, and transmission round times are also acceptable. By implementing variable TD, a trade-off can be made between the number of supported users and transmission round time.

In building a practical system, the round time is not the only factor that needs to be considered, since the waiting in queue time and channel utilization are also critical issues.

Chapter 6

Conclusion

This thesis is focused on access solutions for the media frame network. The current Internet traffic situation is studied first. According to the prediction by Cisco, the worldwide Internet traffic will step into the ZB range by 2015. Among this explosively growing traffic, Internet video is the most significant fraction. A new efficient networking method is needed, especially for larger files.

In the second chapter, a networking methodology was introduced: the media frame network. The fundamental data unit of media frame network is called media frame, which contains typically 10MB of data, transmitted and concatenated as media chains in the network. The most important advantage of MFN is high efficiency and speed for transporting very large files in a reserved end-to-end path. In addition, given the media chain nature, it is anticipated that this network also largely reduces the power consumption of the intermediate routers.

To construct an access network plan for the MFN, the most popular access network standard nowadays is investigated: the passive optical network. By studying the new standards proposed in ITU-T G987 recommendations, we discovered that the physical layer architecture of XGPON could be employed in the access network for MFN. To accommodate the MFN and XGPON, physical layer specifications including wavelength plan, data rate and components logical structure were discussed.

Because of the characteristics of the MFN, it is essential to redesign the protocols over the access physical layer. In chapter 4, the initializing protocols, channel allocation plans and OLT-ONU negotiation protocols were discussed. Different method candidates were investigated as well. To analyze the protocols, a simulation platform based on Omnet++ was built to emulate the access network transporting media chains.

Different transparency degree plans were proposed and graphical results of the system performance were obtained. Fixed TD and variable TD systems were investigated respectively. For fixed TD, transmission is very efficient but the number of supported users can be improved. Variable TD systems make a trade-off between the number of supported users and the transmission efficiency.

Contributions of This Thesis

The idea of the media frame network is new, so a large number of questions remain. For the first time, the methodology of establishing an ability to transport media frames on the current PON standard architecture was investigated. This is a very different way to schedule channel resources compared to the present IP networks.

Secondly, a plausible wavelength allocation approach for the required WDM channel was defined. The wavelength plan based on available channels was discussed, and candidates were selected from those ranges.

Thirdly, time-sharing MAC protocols that are suited for MC/MF upstream transmission were defined. Based on the research of current channel scheduling methods and task schedulers in operating systems, a queuing-request-grant method was proposed. The performance of these MAC protocols was analyzed through simulation under various transparency degrees and bandwidth sharing conditions.

In addition, the tradeoffs among essential factors, such as channel utilization, delay and time to transmit and supported user number were also identified and investigated.

Future Work

This thesis proposed the basic concepts on physical layer and MAC layer of the access network for the MFN, with a simple channel scheduler. Much more work need to be done in the future on higher-level issues and scheduler efficiency.

The channel scheduler in this thesis is built on single FIFO queue. Some modern schedulers typically employ multiple queues to prioritize the incoming requests with QoS (quality of service) configurations.

On higher-level issues, an important debate is focused on the transmission reliability. Whether the sender requires confirmation from the receiver, or the transmission is “best-effort”, still needs to be determined, respectively for media chains and DRFs. Meanwhile, a mechanism needs to be built on handling corrupted media frames. For instance, when the received media chain contains a few corrupted frames, the system needs to figure out how to schedule resending.

Bibliography

- [1] C. Index, *Cisco Visual Networking Index: Forecast and Methodology, 2010-2015 [Visual Networking Index] - Cisco Systems*. 2008.
- [2] V. Cisco, *Entering the Zettabyte Era [Visual Networking Index] - Cisco Systems*.
- [3] B. Whitehead, C.-H. Lung, A. Tapela, and G. Sivarajah, "Experiments of Large File Caching and Comparisons of Caching Algorithms," *Network Computing and Applications, 2008. NCA '08. Seventh IEEE International Symposium on*, pp. 244–248, 2008.
- [4] J. Koomey, *Estimating the Total Power Consumption by Servers in the US and Around the World*. Springer-Verlag New York Inc, 2007, p. 343.
- [5] T. E. Darcie, *Media Frames – Networking With Multi-Mega-Byte Packets*. 2011, pp. 1–7 - To be published.
- [6] R. S. Tucker, R. Parthiban, J. Baliga, K. Hinton, R. W. A. Ayre, and W. V. Sorin, "Evolution of WDM Optical IP Networks: A Cost and Energy Perspective," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 243–252, Feb. 2009.
- [7] T. Battestilli and H. Perros, "An introduction to optical burst switching," *Communications Magazine, IEEE*, vol. 41, no. 8, 2003.
- [8] B. Ganguly and V. Chan, "A scheduled approach to optical flow switching in the ONRAMP optical access network testbed," in *Optical Fiber Communication Conference and Exhibit, 2002. OFC 2002, 2002*, pp. 215–216.
- [9] N. Barakat and T. E. Darcie, "Control-Plane Congestion in OBS Networks," in *Broadband Communications, Networks and Systems, 2006. BROADNETS 2006. 3rd International Conference on, 2006*, pp. 1–3.
- [10] F. Effenberger et al., "An introduction to PON technologies [Topics in Optical Communications]," *Communications Magazine, IEEE*, vol. 45, no. 3, pp. S17–S25, 2007.
- [11] C. F. Lam, *Passive optical networks*. Academic Press, 2007, p. 324.
- [12] G. Kramer, *Ethernet passive optical networks*. McGraw-Hill Professional, 2005, p. 307.
- [13] F. Effenberger, "The XG-PON system: Cost effective 10Gb/s access," *Journal of Lightwave Technology*.

- [14] S. Jain, F. Effenberger, and A. Szabo, "World's first XG-PON field trial," *National Fiber Optic*, Jan. 2010.
- [15] I.-T. S. G. 15, "Recommendation ITU-T G.987.1," pp. 1–52, Jun. 2010.
- [16] F. Effenberger, H. Mukai, and J. Kani, "Next-generation PON-part III: System specifications for XP-PON," *Communications ...*, 2009.
- [17] I.-T. S. G. 15, "ITU-T Rec. G.984.1 (03/2008) Gigabit-capable passive optical networks (GPON): General characteristics," pp. 1–43, Feb. 2009.
- [18] I.-T. S. G. 15, "ITU-T Rec. G.987.2 10-Gigabit-capable passive optical networks: physical media dependent layer specification," Apr. 2010.
- [19] "Metropolitan area network," *Wikipedia.org*.
- [20] V. C. Ravikumar, R. N. Mahapatra, and L. N. Bhuyan, "EaseCAM: An Energy and Storage Efficient TCAM-Based Router Architecture for IP Lookup," *IEEE Transactions on Computers*, vol. 54, no. 5, pp. 521–533, May. 2005.
- [21] T. T. Ye, L. Benini, and G. De Micheli, "Analysis of power consumption on switch fabrics in network routers," in *Design Automation Conference, 2002. Proceedings. 39th*, 2002, pp. 524–529.
- [22] S. Aleksić, "Analysis of Power Consumption in Future High-Capacity Network Nodes," *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 1, no. 3, pp. 245–258, 2009.