Next-Generation Ring Based Self Healing WDM-PON Architecture with Private Networking Capability and Wavelength Sharing

by

Hasan Erkan

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

2008

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

Prof. Mohamed A. Ali

Date

Chair of Examining Committee

Dean Mumtaz Kassir

Date

Executive Officer

Prof. Samir Ahmed, EE Dept, CCNY, CUNY

Prof. Roger Dorsinville, EE Dept, CCNY, CUNY

Prof. Nicholas Madamopoulos, EE Dept, CCNY, CUNY

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

Dr. Andrew Wallace, AT&T

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Next-Generation Ring Based Self Healing WDM-PON Architecture with Private Networking Capability and Wavelength Sharing

By

Hasan Erkan

Adviser: Professor M.A. Ali

There is little argument that today's subscriber access network faces many challenges, and remains a bottleneck in terms of the bandwidth and quality of service it affords to the end-users. While existing broadband access solutions have increased capacity to a few Mb/s for a cable modem or digital subscriber line (DSL) connection, this is still far short of the gigabit line speed necessary to support the evolving rich multimedia and real-time services.

Recent advancements are prompting carriers around the world to consider PON-based Fiber-To-The-Home (FTTH) systems as possible successor to current solutions. FTTH is the ultimate level of access, allowing end users to access the backbone networks through

the gigabit capacity of a fiber optic cable. Among the various PON-based FTTH solutions, single channel Time-Division Multiplexed PON (TDM-PON) and multi-channel Wavelength-Division Multiplexed PON (WDM-PON) architectures are the two most viable candidates. WDM-PON has been positioned as a possible successor and upgrade to TDM-PON-based FTTH implementations.

The transition to FTTH raises a whole range of technical hurdles and complex questions. It is the purpose of this thesis to explore and address some of these challenges and questions, and in the process to capitalize on the key findings to devise a novel, simple and cost effective local access C/WDM-PON-based FTTH architecture that addresses the many challenges on the horizon. The salient features of the proposed architecture is that not only does it addresses the key limitations of today's existing access infrastructure; but also supports most of the crucial features that must be supported by the envisioned next-generation broadband access networks, including simplicity, flexibility, cost-effectiveness, dynamic bandwidth allocation and sharing, guaranteed QoS, private networking capability, and above all survivability.

Specifically, this work proposes and devises a novel self healing ring-based local access C/WDM-PON architecture that efficiently supports dynamic allocation of wavelengths/ timeslots and sharing traffic as well as truly direct private connections among PON end users. The proposed architecture combines the salient features of both traditional static WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and increased security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end users).

Acknowledgements

First and foremost, I sincerely thank my advisor Prof. M.A. Ali. I will always be indebted to him for his constant help and generous support at various stages of my graduate study, and particularly grateful for his endless patience with me. His novel approaches, experience and exceptional ability to find solutions to difficult problems were pivotal in this work.

I also extend my deepest gratitude to Prof. Roger Dorsinville whose insightful and fruitful comments has helped me a lot on the development of the Ring based WDM-PON architecture. I also thank him for serving on my supervisory committee.

I am also grateful to Prof. Nicholas Madamopoulos, Prof. Neophytos Antoniades and Dr. Chris Marty for serving on my supervisory committee. Their insightful recommendations have improved the quality of this work.

I would like to thank my friends Dr. Delowar Hossain, Dr. Murat Cokol, and Ajaz Sana for their insightful comments in many issues in the research.

Finally, and most importantly, I am grateful to my mother, Hatice Dogdu for her unprecedented struggle to bring me up to these days, and my wife, Eylem Erkan, for her continuous patience and support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1. Overview

There is little argument that today’s subscriber access network, shown in figure 1.1, faces many challenges, and remains a bottleneck in terms of the bandwidth and quality of service it affords to the end-users. The change in the fundamental character of access networks traffic, as demonstrated by the current shift from voice and text-based services to a bundle of bandwidth hungry converged services that include IPTV, VOIP, gaming, video streaming, and peer-to-peer applications, is further highlighting the gravity of this problem. This has thus seriously impeded the growth of truly broadband services and applications. Once the bandwidth and quality barrier of today’s access networks is removed, new and unforeseen applications will emerge and attain widespread popularity.

While existing broadband access solutions have increased capacity from the range of 56 kb/s for a dial-up modem to a few Mb/s for a cable modem or digital subscriber line (DSL) connection, this is still far short of the gigabit line speed necessary to support the evolving rich multimedia and real-time services. There is a growing perception that copper access networks will soon no longer be able to meet the ever-growing consumer demand for bandwidth. This, along with a combination of regulatory and competitive forces, as well as recent rapid advances and standardizations of Passive Optical Network (PON) technology [1-5], are finally prompting carriers around the world to consider PON-based Fiber-To-The-Home (FTTH) systems as possible successor to current copper-based access solutions. FTTH is the ultimate level of access, allowing end users to access the backbone networks through the gigabit capacity of a fiber optic cable.

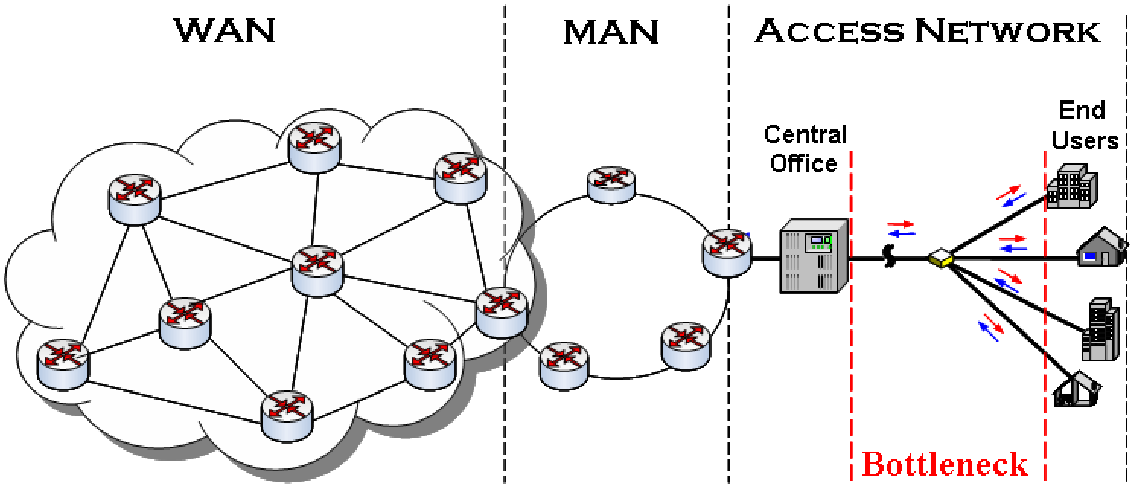


Figure 1.1 Access Network Bottleneck

A PON connects a group of Optical Network Units (ONUs) located at the subscriber premises to an Optical Line Terminal (OLT) located at the service provider's facility. It consists of a single, shared optical fiber (trunk) connecting an OLT to a passive star coupler (SC), which splits the downstream signal to multiple ONUs over dedicated short optical fiber. Traffic from an OLT to an ONU is called "downstream" (point-to-multipoint), and traffic from an ONU to OLT is called ‘upstream’ (multipoint-to-point). Among the various PON-based FTTH solutions, single channel Time-Division Multiplexed PON (TDM-PON) and multi-channel Wavelength-Division Multiplexed PON (WDM-PON) architectures are the two most viable candidates.

Due to their reduced operational and equipment costs, TDM-PONs have been widely accepted as a viable technology for the implementations of FTTH solutions, and are being deployed in the field in several places around the world. TDM-PON-based FTTH access solutions including broadband PON (BPON), gigabit PON (GPON), and Ethernet PON (EPON) are finally emerging into the mainstream and are set to revolutionize the access infrastructure worldwide. It is widely anticipated that, over the next decade, copper access networks worldwide will be largely replaced by fiber access networks, marking the beginnings of a new era of a mass migration to PON-based FTTH solutions.

1.2. Thesis Motivation

TDM-PON supports a single wavelength channel in the downstream direction (OLT to ONUs) and another wavelength in the upstream direction (ONUs to OLT). Current TDM-PONs including BPON, GPON, and EPON provide higher bandwidth than traditional copper-based access network solutions and seem to offer a satisfactory solution for present bandwidth demands. However, since all users attached to a TDM-PON share a single transmission channel, the average dedicated bandwidth assigned to each user in either direction is usually limited to a few percent of the channel capacity, i.e., a few tens of Mb/s. But as the number of end users attached to a given access network grows each year, and as their usage patterns evolve to include more bandwidth-intensive applications, there emerges an acute need for future broadband access solutions that can upgrade current TDM-PONs and provide each user with at least a 100 Mb/s, symmetric, and guaranteed bandwidth [5].

WDM-PONs are emerging as the most promising next-generation access solution that can provide evolutionary upgrade to existing TDM-PONs. WDM-PON has been positioned as a possible successor to TDM-PON-based FTTH implementations. Though technically feasible, compared to TDM-PON, it is still considered too expensive for widespread residential deployment. However, as we argue in this work, this is only true for low-bandwidth applications, where utilizing TDM-PON might be a little less expensive, but as bandwidth demand increases, the economics change. "In terms of cost per bit rate," WDM-PON is more efficient and economical." To date, Korea Telecom (KT) has been the most aggressive proponent of WDM-PON, investing $3 million in a WDM-PON-based FTTH deployment in Gwangju.

Traditional WDM-PON systems allocate a separate pair of dedicated upstream and downstream wavelength channels to each subscriber, enabling the delivery of a symmetric 100 Mb/s or more of dedicated bandwidth per subscriber/ONU in each direction. In contrast to traditional TDM-PON architectures, WDM-PON has a unique ability to support any future bandwidth upgrade without altering the physical infrastructure. In addition to its operational simplicity, WDM-PONs provide dedicated point-to-point optical connectivity to each subscriber with bit rate and protocol transparencies (can support the transmission of any service or service mix, including Gigabit Ethernet, ESCON, FICON, Fiber Channel, OC-N, ATM, and FDDI, at any bit rate), guaranteed QoS, and increased security.

Despite these numerous crucial advantages, traditional WDM-PON architectures suffer from several limitations including inability to efficiently utilize network resources and to cope with the dynamic and bursty traffic patterns of the emerging integrated triple pay services [5-6]. This limitation is more pronounced under non-uniform network traffic loads scenario, i. e., when some subscriber's wavelength channels, for a given interval, are heavily loaded while others are underutilized or are totally idle. In this case, the unused dedicated channel capacities of those lightly loaded/idle subscribers cannot be shared by any of the other heavily loaded users attached to the PON, leading to the waste of scarce network resources. Note that this traffic scenario is the norm in most real practical networks traffic, particularly when a significant fraction of these networks traffic is dynamic and bursty (fluctuates significantly with time). In other words, this limitation would have more significant impact on the performance of the envisioned next generation broadband access networks, since the dominant traffic of these networks (triple play services) is dynamic and bursty. Therefore, it is essential that future WDM-PON architectures must support dynamic bandwidth allocation (DBA) and sharing.

To address this problem, several WDM-PON architectures and protocols that dynamically manage and allocate bandwidth in both time and wavelength dimensions have been proposed recently [6-11]. Most of these schemes, however, are costly and assume complex OLT and ONUs setups, which require tunable, or array of fixed transceivers, or both, WDM filters, and wavelength-band-selective receivers. Furthermore, schemes that support dynamic wavelength sharing [6-7], where additional wavelength channels are added to accommodate the fraction of bursty downstream traffic that may exceed the user's dedicated downstream wavelength channel rate, are still falling short of addressing the fundamental problem of the inefficient utilization of network resources. This is because the unused capacities of those lightly loaded/idle dedicated downstream wavelength channels are still being wasted.

In addition, TDM/WDM-PON architectures are traditionally deployed as tree topologies, which provide added flexibility and simplicity in the deployment of access networks. However, tree-based topologies have several inherent limitations including:

1. Inability to support a truly shared Local Area Network (LAN) capability among end users.
2. Lack of a simple and cost-effective protection and/or restoration capabilities, specifically against failures in the distribution network for both TDM and WDM-PONs. ITU-T G.983.1 recommended four possible TDM-PONS protection schemes, which duplicate fibers and equipment at the ONUs and OLT [6]. There are no available WDM-PON Standards yet. However, applying these redundant protection schemes to WDM-PONS would be cost prohibitive.
3. To date, mainstream PON DBA and protection schemes have been centralized– relying on a component at the distant OLT to arbitrate upstream transmission and to detect and recover distribution and trunk fiber breaks [1-6, 8-11]. In addition to

the typical “single-point of failure” problem (failure of OLT software will bring down the whole access network), the centralized processes of upstream bandwidth allocation as well as detecting and restoring distribution fiber breaks at the distant OLT are lengthy and complex processes and require many changes at each ONU [7].

Deploying point-to-point customer communication links to emulate a shared local area network (LAN) within a single PON infrastructure is an important feature for providing a private networking capability, which has recently received some attention [12-16]. In general, standard upper layer shared LAN emulation techniques require the use of bridges/routers at the OLT to redirect data frames back to the ONUs. These techniques are effective but increase the cost and complexity of the network and reduce both available downstream and upstream transmission bandwidth. Several physical layer LAN emulation techniques have been proposed to achieve intercommunication among ONUs, but only within the context of TDM-PONs [12-14].

Most recently, fewer schemes have been proposed to achieve intercommunication among the ONUs within a WDM-PON infrastructure [15-16]. In addition to the added cost and complexity, most of these schemes, however, suffer from poor scalability due to high splitting losses as the redirected LAN signals traverse through the star coupler/AWG once or twice, resulting in lower power budget that limits the number of ONUs that can be attached to a single PON. In general, achieving intercommunication among subscribers within a tree-based WDM-PON setup is a lengthy and complex process that requires much more resources than those needed for a TDM-PON.

On the other hand, typical protection scenario adopted for WDM-PONs utilize duplicated network resources such as fiber links and/or equipment at the ONUs, along with automatic protection switching (APS) to reroute the affected data traffic into those alternate predetermined protection paths. Several survivable WDM-PON architectures [5-10] including a group protection scheme [5-7] and an alternate path switching scheme [9] have been proposed to protect against both the feeder and distribution fibers. In most of these schemes, the periodic and cyclic properties of arrayed waveguide gratings (AWGs) are used to enable each distribution fiber to carry data traffic for more than one ONU [5-7]. The underlying strategy is to inter-connect two adjacent ONUs [5-7] or RNs [8] by a piece or a pair of fiber such that the affected bi-directional traffic could be re-routed via the adjacent ONU or RN.

The work in [9] re-routes the data traffic to the affected ONU by utilizing a second AWG at the RN along with a second distribution fiber that inter-connects the second AWG with the affected ONU. Protection switches are incorporated in either each ONU [5-8] or in each transceiver located at the OLT [9]. These schemes are all centralized, complex, and can significantly alter the cost-effectiveness of PONs since they require many redundant components (multiple AWGs, red/blue passband filters, optical switches, optical isolators, Erbium-doped fiber amplifiers (EDFAs) at the OLT, and many spare fiber connections to each ONU) as well as wavelength assignment schemes.

Though these numerous aforementioned WDM-PON architectures [5-16] have tackled the limitations of traditional WDM-PONs, however, each of these proposed architectures has only addressed just a single limitation. All of these architectures apply point-solutions and work-arounds to the existing traditional WDM-PON. The fact remains that these point-solutions, which incrementally applied to the traditional WDM-PONs result in increased cost and complexity. The end result is that although these architectures are technically feasible, they are still considered too complex and expensive for widespread near-term deployment of WDM-PONs. Event if these WDM-PON architectures were to take hold in the future, it is likely to be a rather distant future. For WDM-PON-based FTTH solutions to evolve as the envisioned next generation broadband access networks, at a minimum, they must support all of the features offered by existing TDM-PONs, including simplicity, flexibility, cost-effectiveness, dynamic bandwidth allocation, and private networking capability.

1.3. Thesis Statement

The transition to FTTH raises a whole range of technical hurdles and complex questions. It is the purpose of this thesis to explore and address some of these challenges and questions, and in the process to capitalize on the key findings to devise a novel, simple and cost effective local access C/WDM-PON-based FTTH architecture that addresses the many challenges on the horizon. The proposed architecture builds upon a new direction to the design and implementation of next generation broadband access networks, and so is likely disruptive. The salient features of the proposed architecture is that not only does it addresses the key limitations of today's existing access infrastructure; but also supports most of the crucial features that must be supported by the envisioned next-generation broadband access networks, including simplicity, flexibility, cost-effectiveness, dynamic bandwidth allocation and sharing, guaranteed QoS, private networking capability, and above all survivability.

Some of the challenging questions as well as the key debates in the FTTH market that this work will address are: (1) Can the high cost of FTTH deployment really be justified? (2) Which architectures and technologies best meet requirements? (3) Whether future FTTH will be dominated by TDM-PONS or WDM-PONS? And whether it will be built on open-access principles – available on a wholesale basis to all service providers – or as a closed network, with access restricted to or controlled by the builder.

This work examines the technological requirements and assesses the performance analysis and feasibility for implementing a novel WDM-PON-based FTTH broadband access architecture that addresses the key limitations of conventional tree-based PON architectures including: a) providing a simple and cost-effective fully distributed/hybrid resilience capabilities against most types of networking failures; b) supporting private networking capability as well as dynamic allocation and sharing of network resources. Specifically, this work proposes and devises a novel self healing ring-based local access C/WDM-PON architecture that efficiently supports dynamic allocation of wavelengths/ timeslots and sharing traffic as well as truly direct private connections among PON end users. The proposed architecture combines the salient features of both traditional static

WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and increased security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end users).

Unlike a typical WDM metro-access ring network, where the feeder fiber of a PON is replaced with a metro fiber ring that interconnects the hub and access nodes, the proposed architecture interconnects WDM ONUs via a short distribution fiber ring in the local loop but allows them to share the feeder fiber for long reach connectivity to the OLT. This architecture is well suited for an autonomous access environment such as a private financial enterprise where supporting secure, reliable, and direct intercommunications among end users is required and where several high-end users are closely dispersed within a 1-10 km diameter area.

To achieve our overall objectives, this work will be divided into four main phases. In the first phase, the basic building blocks, components and sub-systems required to build both the normal and survivable states of the propose FTTH access architecture, along with their associated functionalities, pros and cons, as well as their impact on the overall system performance are analyzed and presented. Specifically, a detailed analysis of the performance impact of both downstream and upstream power budgets along with their impact on the overall performance and scalability of the proposed architecture is presented.

In the second phase, a fully distributed control plane, which achieves intercommunication among the ONUs is devised and presented. The control plane is shown capable of supporting fully distributed as well as hybrid (distributed/centralized) fault detection and recovery schemes. Supported by the distributed control plane, we develop a decentralized LAN DBA scheme that dynamically allocate network resources among end users fairly and efficiently, in which the OLT is excluded from both bandwidth arbitration and fault detection/recovery processes.

In the third phase, a resource allocation scheme that efficiently supports dynamic allocation of both wavelengths and timeslots and sharing traffic among PON end-users is developed. Specifically, we develop an OLT-based shared wavelength/timeslots assignment and scheduling (SWS) algorithm that dynamically allocate network resources among end users fairly and efficiently. To achieve this objective, overall network resources, including downstream, upstream, and LAN wavelength channels, are load-balanced and efficiently utilized via traffic-engineered routing of subscriber's traffic. The algorithm is QoS-aware via assigning dedicated wavelength channels for higher priority traffic (e. g., voice and video) and limiting dynamic wavelength assignment and sharing to best-efforts traffic only.

Finally, the fourth phase presents and devises fully distributed as well as hybrid (distributed/centralized) cost-effective and efficient fault detection and recovery schemes. The performance as well as the pros and cons of each scheme are compared with each other and with those of traditional centralized APS protection schemes. The proposed protection schemes are capable of protecting against both node and distribution/trunk fiber failures. These schemes enable the restoration of all network traffic including upstream, downstream, and LAN data. In addition, these schemes can also protect against any combination of concurrent double failures including trunk/distribution fiber breaks and node failures.

Furthermore, the recovery time associated with any and all different distribution network/trunk failures is still within the delay-bound limit required for delivering guaranteed triple play services. In addition to the added flexibility and reliability of a distributed scheme, the proposed architecture eliminates the OLT's centralized task of failure detection and subsequent recovery scenarios as well as the task of processing requests and generating grants (GATE messages) for bandwidth allocations. This reduces the additional processing complexities and delays at the OLT.

It is shown that the proposed distributed solutions can overcome the problems associated with current centralized ones, and in the process prove that these distributed networking architectures and the associated fault detection and recovery schemes as well as bandwidth allocation algorithms and protocols have characteristics that make them far better suited for supporting cost effective resilience capabilities and for provisioning QoS schemes necessary for properly handling data, voice, video, and other real-time streaming advanced multimedia services over a single line.

1.4. Organization of Thesis

Organizational outline of the thesis is as follows:

Chapter 2 describes the history of access solutions. The transition from cable modem (CM) and digital subscriber line (DSL) to standard single channel PON systems and eventually to WDM-PON system is discussed. Finally an overview of WDM-PON systems is given.

Chapter 3 investigates the viability of implementing fully distributed control plane architecture for EPON system. The details of distributed architecture along with its operational principles are described. In the proposed distributed architecture, ONU intercommunication is achieved without OLT intervention (OLT is relieved from that burden). Our setup shows that with a 1dBm transmitted power/ONU along with –38 dBm measured receiver sensitivity; a power budget of 39 dB is adequate to support 32 ONUs. This number could be further increased with higher transmitter power and receiver sensitivity.

Chapter 4 proposes a cost effective open access CWDM PON scheme. Government regulations in many countries have mandated a clear demarcation between the network operators who provide physical connectivity and transport data (e.g., telephone and cable companies), and the service providers (SPs) who deliver the content and the services (e.g., Internet SPs). The primary motivation is to have free market competition between SPs, thereby making services cheap and available to end users; and also preventing monopoly by the network operator who owns the right-of-way to lay cable/fiber in a residential area. Such a framework is known as ***open access.*** We believe open access is achieved best with the proposed CWDM-PON architecture since network operators assigns a set of wavelengths to each SPs. Since each SP operates on different set of wavelengths, their services are transparent to each other.

Chapter 5 proposes a simple and cost effective Ring-Based Local Access C/DWDM-PON Architecture which supports shared LAN capability among end-users.

Chapter 6 discusses the importance of survivability of ring-based WDM-PON architecture. Two scheme (hybrid and distributed) is proposed for the proposed WDM-PON Ring architecture. An APS protection scheme is introduced. A comparative analysis between two schemes is investigated.

Chapter 7 gives concluding remarks and future challenges in this work.

Chapter 2

Overview of Access Solutions

2.1 Legacy Access Technologies (Cable Modem and Digital Subscriber Line)

Until recently, two of the most popular technologies that offer speedy access to the World Wide Web have been Digital Subscriber Line (DSL) and the cable modem. There are several reasons why this is the case and the most important one is that both of these internet connections are considerably faster than the standard dialup connections

A cable modem is an external device that allows one’s computer to connect to the Internet through a cable TV wire (coaxial cable), instead of a telephone line (twisted twisted-pair copper). On the operator side (Head-End.), user’s cable is terminated by Cable Modem Termination System (CMTS) at the Cable TV head end.

Due to high loss at the coaxial cable, in practical applications, electrical amplifiers are used every 100 to 200 m with gains of about 20 dB. The number of amplifiers is limited by the required signal-to-noise ratio (SNR) to deliver an acceptable quality video signal.

This places practical limits on the transmission distance using only coax cable. To overcome this limitation, single-mode fiber is used to feed the broadcast analog-video signals (50–100 channels) to a fiber node (FN). They are then converted to electrical-video signals for transmission along multiple-coaxial cables. These final transmission links use periodically amplified coax-cable networks that typically consist of tree-and-branch architectures [1]. This combination of fiber and coax cables is called a hybrid fiber-coax (HFC) network. Each FN typically covers a “cell” of about 500–1000 homes. However, since 500–1000 subscribers share the bandwidth within a cell, the guaranteed bandwidth per subscriber is only 2.8–5.6 Mb/s for the downstream signal and 0.15–0.3 Mb/s for the upstream signal. HFC architecture is also known as FTTN/Cab fiber to the node/cabinet) architecture, due to utilization of optical fiber between Cable TV head-end and FN.

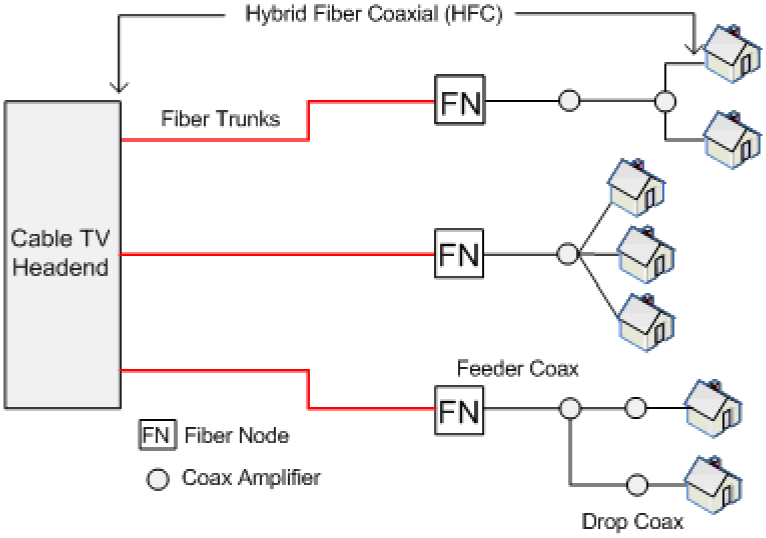


Figure 2.1: CATV architecture

The most widely deployed transmission media in access networks are twisted-pair copper cables that have been used for telephone lines for more than 100 years.

DSL uses the same twisted pair as telephone lines and requires a DSL modem at the customer premises and Digital Subscriber Line Access Multiplexer (DSLAM) in the central office (CO). The most common DSL technology is Asymmetric Digital Subscriber Line (ADSL). In ADSL, the overall downstream rate depends on, the distance covered, the size of the wire, and interference. The main drawback of DSL systems is the limited transmission distances. DSL have the maximum distance limitation of 18000 feet (~5.5km). A figure of merit for the bit-rate-times-distance product of twisted-pair copper cable is about 10 Mb/s · km. This means that a 100-Mb/s signal can be transmitted only over a distance of about 100 m, which greatly limits its use in high-speed access networks.

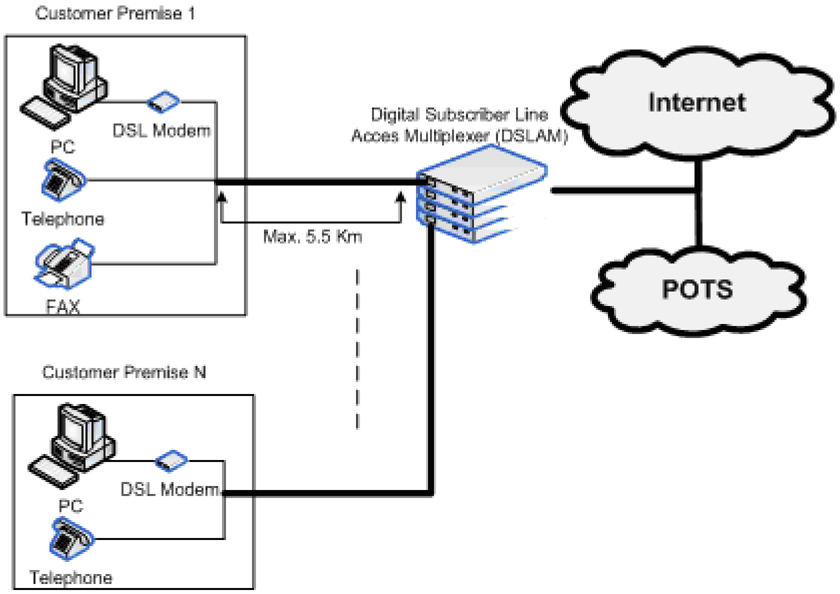


Figure 2.2: ADSL architecture

Very High-speed DSL (VDSL) is proposed as an evolutionary upgrade to ADSL. VDSL requires the utilization of fiber between CO and street cabinet and twisted copper from street cabinet to the homes. VDSL, like its predecessor ADSL, is also distance limited technology where copper length should not exceed 300m to achieve promised data rate (52 Mbit/s downstream and 6.4 Mbit/s upstream). VDSL architecture is also known as FTTC (Fiber to the curb) architecture since optical fiber is used between CO and street cabinets.

Even though both DSL and Cable Modem (CM) technologies offer some improvement, but they still do not provide enough bandwidth for emerging integrated triple play services and applications such as Video-On-Demand (VoD), telemedicine or two-way video conferencing.

2.2 FTTx Solutions on the Access Network

Optical fibers are penetrating deeper into the first mile. The most promising solution to the first mile bottleneck is to bring the fiber all the way to the doorstep of the customers. This solution is called as Fiber-to-the-home/premise (FTTH/P) or Fiber to the Building (FTTB). Through FTTH/P, the backbone capacity (Gb/s or even Tb/s) is carried all the way to the customer premises/home solving the access bottleneck problem once and for all. Other variations of FTTx are fiber to the node (FTTN) and fiber to-the-curb (FTTC). Since FTTN and FTTC require usage of metallic cables at the local loop, these two technologies are not viewed as the viable solution for the first mile bottleneck problem. Schematic diagram for all FTTx technologies is given in Figure 2.3.

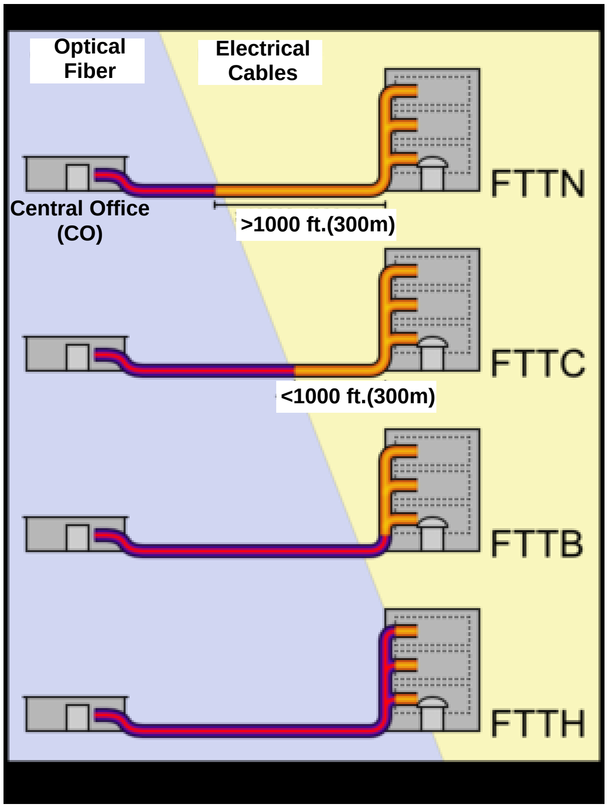


Figure 2.3. A schematic diagram of FTTX architectures

2.2.1 FTTH Deployment Scenarios

Potential FTTH deployment scenarios are given in figure 2.4. Since access networks mainly targets residential users as potential customers, to be commercially viable, they have to be cost effective and simple. As an illustrative example of 32 end-users, Point-to-Point (PtP) solution requires 32 fibers and 64 transceivers. The huge fiber deployment comes with the labor cost which makes this solution expensive for near term deployment. Second solution is called Curb-Switched FTTH or Active Optical Network.

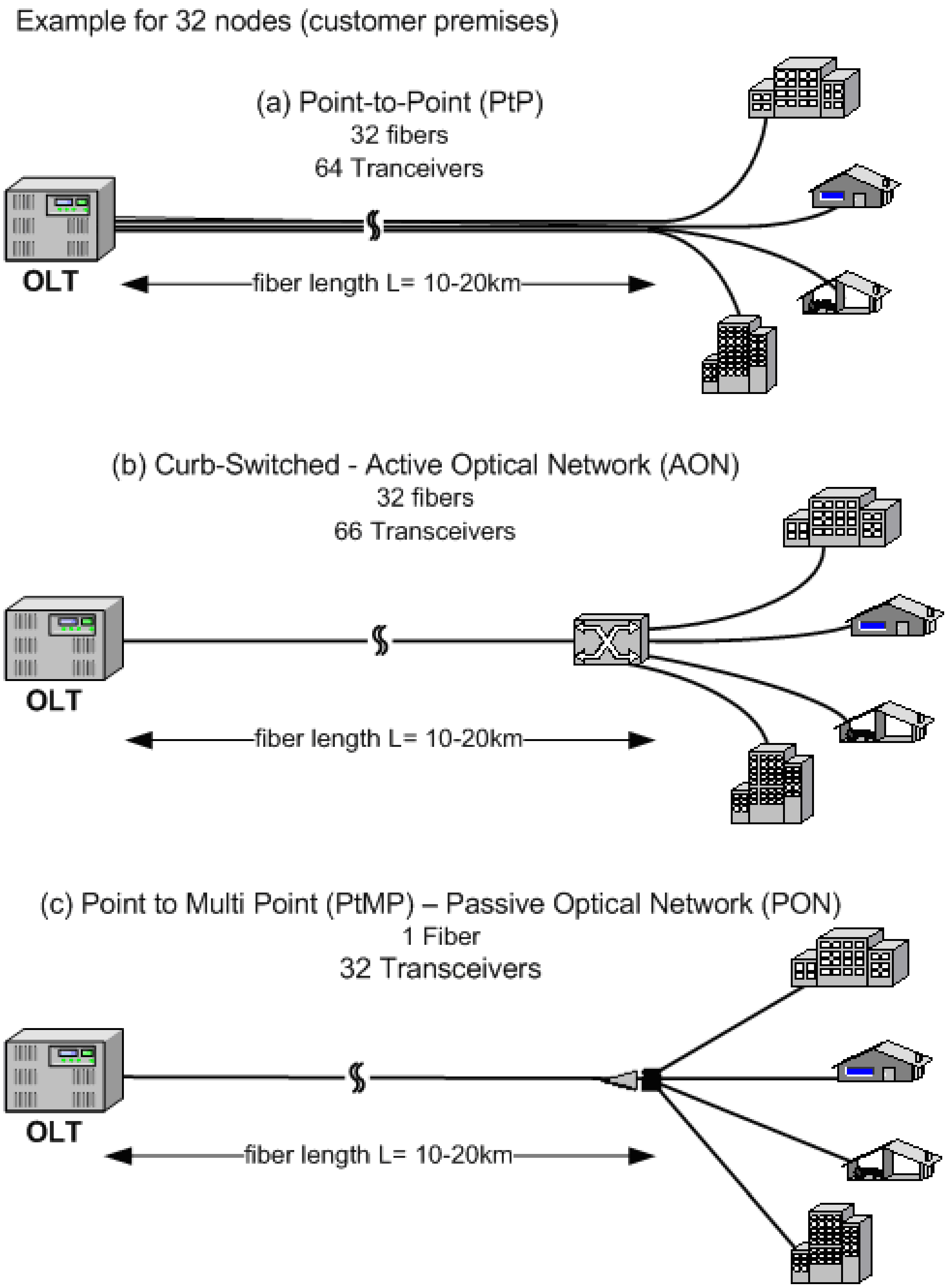


Figure 2.4 FTTH architectures

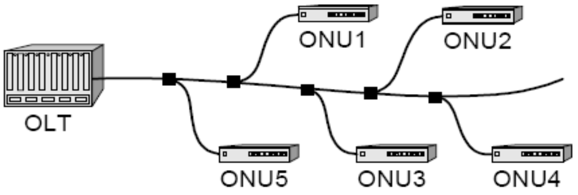
Telecommunications companies are rather hesitant when it comes to applying active equipment in the local distribution network. It is intrinsically more error-prone, more sensitive to changes in temperature and sensitive to radio interference. An especially important factor is that active equipment introduces the additional risk of down time due to local power failures. Additionally, AON networks comprise complex and expensive active devices such as Add-Drop Multiplexers (ADMs), Digital Cross Connects (DCS), Regenerators (repeaters), etc. These networks bring unlimited bandwidth to support large number of users over a vast area. The complexity and cost of AON is feasible in MAN, where numerous users share the expense. Compared to MANs, subscriber access networks serve a relatively small number of users within short distance, requiring much less bandwidth. Therefore, access network is very cost sensitive. Due to the mentioned technical complexity and higher cost of AON, current FTTH systems are not commercially viable; thus have prevented the deployment.

Third scenario for FTTH deployment is called as Point to Multi Point (PtMP) FTTH or Passive Optical Network (PON). This solution is not only most simple but also most cost effective scenario among proposed three different scenarios. Passive Optical Network (PON) eliminates active electronic component of AON, such as amplifiers, regenerators, from the outside plants and replace them with passive, inexpensive, long lasting, easy to maintain optical splitters & couplers. Following section gives detailed explanation about PONs.

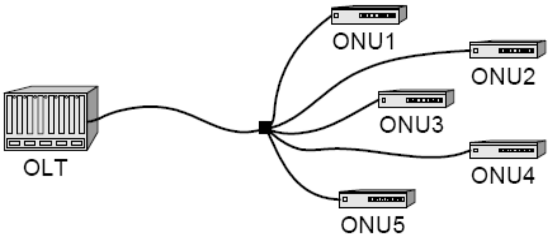
2.3 Passive Optical Networks (PONs)

A Passive Optical Network (PON) is a group of technologies originally created by the Full Service Access Network (FSAN) working group and now standards of ITU-T and IEEE, allowing fiber to reach all the way to the customer premises (FTTP). A PON consists of a central office node Optical Line Termination (OLT) at the service providers' office and a number of Optical Network Units (ONUs) near end users, and the fibers and splitters between them, called the optical distribution network (ODN). The OLT provides the interface between the PON and the backbone network, while the ONT provides the service interface to the end user. PON is a converged infrastructure that can carry triple play services over the PON fiber.

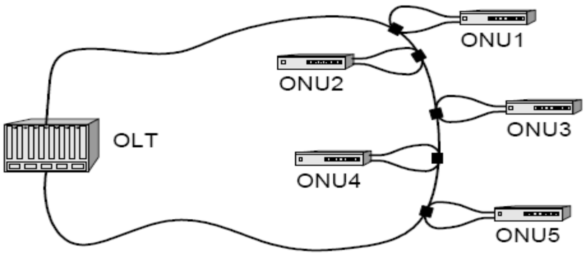
PON does not have any active elements in the signal's path from central office to the customer-premises equipment. Unlike previous architectures, where fiber is used as a feeder to shorten the lengths of copper and coaxial networks, PON deployments use optical fiber throughout the access network, reaching all the way to customer premises. PON technologies are capable of supporting gigabit per second (Gbps) speeds, at costs comparable to those of current solutions. Three proposed PON deployment scenarios are tree architecture, bus architecture and ring architecture. These architectures are shown in the following figure.



(a) Tree



(b) Bus



(c) Ring

Figure 2.5 PON architectures

2.3.1 PON Technologies

The major PON technologies under consideration by service providers as the basis for the FTTH deployments are: Broadband PON (BPON), Ethernet PON (EPON), and Gigabit PON (GPON) [2-11]. Transmission standards utilized in PON networks are based on ATM and Ethernet technologies. BPON standards are specified in ITU G.983, EPON standards are specified in IEEE 802.3ah, and GPON standards are specified in ITU G.984.

Earlier PON deployments utilized ATM PON (APON), are evolved into broadband PON (BPON). The A/BPON protocol is characterized by having two downstream wavelengths and one upstream wavelength. The 1550 nanometer (nm) and 1490 nm wavelengths are used for downstream traffic, with the 1490nm channel typically an IP channel for voice and data service. The 1550nm channel will be used for a radio frequency (RF) or IP video overlay. Providing 622 Mbps shared electronics are able to dynamically provide 20 to 30 Mbps per subscriber. Time division multiple access (TDMA), recommended by FSAN, is used for all down/upstream traffic.

An alternative to A/BPON networks is Ethernet PON (EPON). EPON only uses two wavelengths and exclusively uses IP. The 1550 nm wavelength is used for downstream traffic, and the 1310 nm wavelength is used for upstream traffic. Capable of 1.25 Gbps in shared bandwidth, EPON under "best-effort" conditions provides for 100 Mbps but typically provides for bandwidth of 30 to 40 Mbps.

Broadband PON has evolved into Gigabit PON (GPON) to address bandwidth and protocol limitations. Capable of up to 2.5 Gbps shared bandwidth among 32 users; GPON utilizes the same wavelength plan of BPON. It is governed under ITU standard G.984 and provides for protocol flexibility across ATM, Ethernet, and TDM platforms.

Ethernet currently accounts for more than 85 percent of all installed connections. As opposed to the ATM networks, since Ethernet networks are simple and inexpensive, it has become a universally accepted standard, with over 320 million ports deployed worldwide [12].

One of the most important factor in analyzing PON solution's overall cost is the efficiency factor, providing the overall bandwidth that can be sold as services over the system. This bandwidth is called revenue bits. Among the PON solutions, GPON is the best choice in terms of revenue bits while EPON is the worst [13].

Table 2.1 provides a complete breakdown of the PON protocols and the respective capabilities

Features	BPON	EPON	GPON
Standard Body	ITU-T G983	IEEE 803.2ah	ITU-T G.984
Max. Speed (Upstream)	622 Mbps	1.25 Gbps	2.5 Gbps
Max. Speed (Downstream)	1.25Gbps	1.25 Gbps	2.5 Gbps
Number of ONTs	16/32/64	16/ 32	16/32/64/128
Downstream Capacity	520Mb/s(for 622Mb/s)	910Mb/s	1170Mb/s(for 1.244 Gb/s)
Upstream Capacity	500Mb/s(for 622Mb/s)	760-860 Mb/s	11600Mb/s(for 1.244 Gb/s)
Payload	ATM	Ethernet	ATM, Ethernet, TDM
Logical reach (km)	20	20	60*
Cost	High	Low	High
Complexity	High	Low	Undetermined
Network Interfaces	ATM, TDM, GE	GE	ATM, TDM, GE
Access Interfaces	ATM, TDM, 10/100	10/100	ATM, TDM, 10/100
Downstream λ (nm)	1490 and 1550	1550	1490 and 1550
Upstream λ (nm)	1310	1310	1310
Line coding	NRZ (+ scrambling)	8B/10B	NRZ (+ scrambling)
Downstream Security	AES	Not defined	AES

Table 2.1 PON Protocols

2.3.2 Multiplexing Methods in PON Networks

All point to multipoint systems requires a media access scheme specifically in upstream direction to avoid collision and share the upstream channel [14]. These schemes can be generally categorized as time-division multiple access (TDMA), wavelength division multiple access (WDMA), subcarrier-division multiple access (SCMA), and code-division multiple access (CDMA). Due to technological limitations of latter two, the former two architectures received more attention and most likely those schemes will see widespread future use. Hence, next two sections discuss TDMA and WDMA architectures only.

2.3.2.1 Time-Division Multiple Access (TDMA)

TDMA is one of the most popular techniques being considered for PON upstream access among end users. In this scheme, OLT assigns a dedicated time slot to each of the multiple subscribers connected to the PON. Due to sharing, each subscriber can use the full upstream bandwidth of the optical link for the duration of its allotted time slot. The average dedicated upstream bandwidth for each subscriber is usually only a few percent of the channel capacity.

In this scheme, OLT assign time slots to each user to avoid any collisions of the data in the upstream direction. ONTs need to negotiate with the OLT of when it can send its data. To realize an efficient use of the shared communication channel, the OLT should have precise time-delay information to each ONT since they are all at different distances from the CO. This information is needed to control the time when each ONT can transmit to ensure that no packets overlap when they reach to the remote node (RN). The process to identify the actual delay time to each ONT is known as ranging or discovery.

When TDMA is used for upstream access, then the system is called as TDM-PON. TDM-PON schemes require a burst-mode optical receiver at the OLT and burst mode optical transmitter at the ONTs. These OLT receivers can be technically challenging since they should have a wide dynamic range for the different incoming burst signals and they require a very short guard time for the clock-synchronization process. In addition, each ONT transmitter must be sure to completely shut down after its time slot is over since any residual background light will interfere with the upstream data signals from the other ONTs. Following figure shows a typical upstream access for a PON scheme.

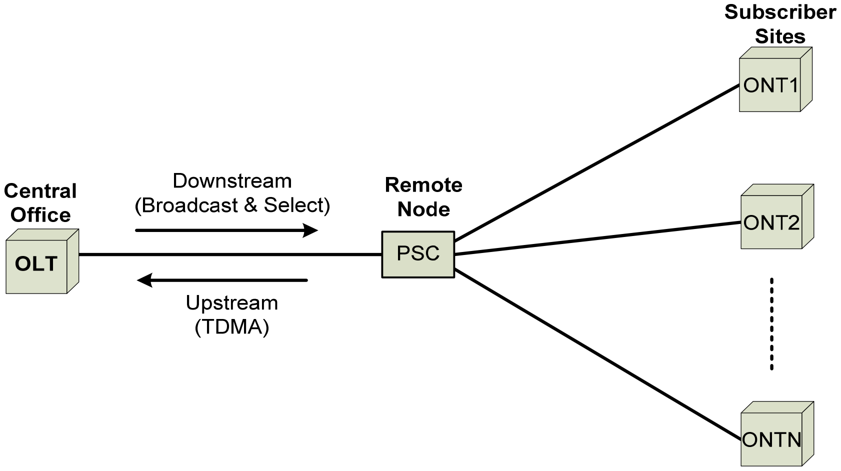


Figure 2.6 TDM-PON architecture

2.3.2.2 Wavelength-Division Multiple Access (WDMA)

In WDMA scheme, each subscriber is assigned a pair of dedicated wavelengths which are typically used for user's upstream and downstream traffic transmission. There is no interaction or coupling between the subscribers on a WDM-PON. That is why; each user can send data to the OLT at any time, independent of other users. This eliminates any management issues related to sharing the up-/downstream wavelengths that is typically seen in TDM-PON.

In WDM-PON, there are no requirements for burst-mode receivers or transmitters and no need for any sophisticate MAC algorithms to manage the timing of the ONT transmissions. This makes the operation of a WDMA network very simple. In addition to its operational simplicity, WDM-PON approach provides dedicated point-to-point optical connectivity to each subscriber with very high bit rate. Other advantages of WDM-PON technologies over TDM-PON are; protocol transparencies among end-users, guaranteed QoS for each end-user, and very high security which is typically a big concern for broadcast and select systems.

To realize WDMA functionality, WDM multiplexer/demultiplexer (MUX/DEMUX) is used at the OLT and RN. To multiplex and de-multiplex separate wavelengths, traditionally, a variety of technologies are utilized such as Thin Film Filter (TFF) based MUX/DEMUX, Arrayed Waveguide Grating (AWG) based MUX/DEMUX, or Fiber Brag Grating based MUX/DEMUX. Especially, when AWG MUX/DEMUX is used the insertion loss at the RN is considerably smaller and effectively independent of the splitting ratio. Hence, the number of end-user connected to a WDM-PON system is much more than TDM-PON systems which are typically limited with power budget due to high splitting ratio at RN.

On the other hand, WDMA architectures are still cost prohibitive when compared to TDMA systems due to the number of transceivers required is almost double of TDMA architectures (2N versus N+1). Research is underway to find the cost effective solution for WDM-PON systems since cost is the key bottleneck in the commercialization of WDMA networks. Another potential concern is that the peak data rate to each user in both downstream and upstream is equal to their dedicated data rate. In other words, the idle or underutilized wavelength channels both in downstream and upstream cannot be shared by congested wavelengths. Considering the bursty nature of the internet traffic, this concern is exacerbated. Following figure shows a typical WDM-PON architecture.

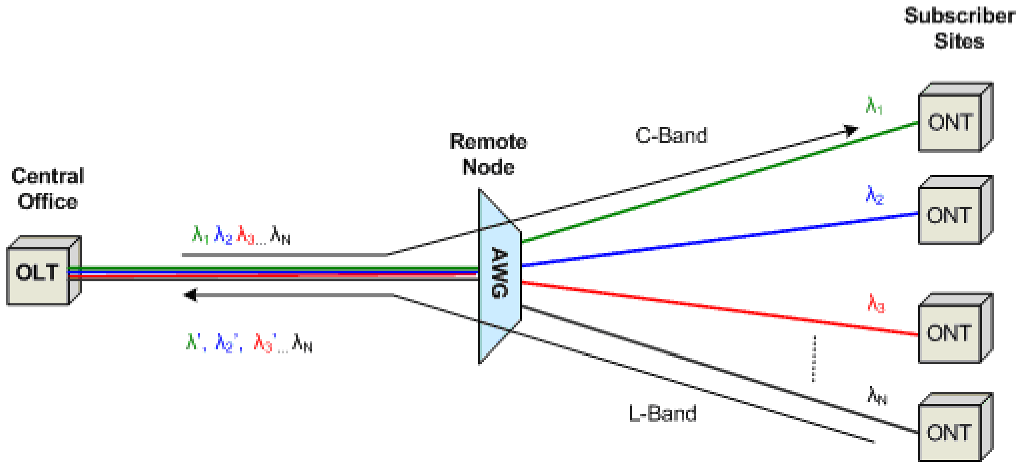


Figure 2.7 WDM-PON architecture

2.4 Conclusion

This chapter discusses that due to recent advances in optical networking technology, backbone (WAN and MAN) capacity reached to Tb/s speeds. But due to technological limitations of current access schemes (xDSL and Cable Modem), this capacity is not reachable by the end users. This problem is called last mile bottleneck or access bottleneck. To face the challenge of access bottleneck, optical fibers are penetrating deeper into the first mile. FTTH is the ultimate level of access, allowing end users to access the backbone networks through the gigabit capacity of a fiber optic cable. Among proposed three schemes for FTTH, PON drew the attention of network community as the ultimate access solution, due to its efficiency, simplicity, and economical feasibility.

PONs are classified into different categories based on the data link layer protocol utilized. Those are namely are BPON (APON), EPON, and GPON. PONs can also be categorized based on multiplexing technique used in upstream access.

The most promising multiplexing techniques used in PON schemes are TDMA and WDMA. If TDMA used in the PON scheme, then the system is called as TDM-PON and if WDM used in the PON scheme, then it is called as WDM-PON. WDM PON scheme is considered as upgrade version of TDM PON scheme where each subscriber (ONT) is assigned a separate pair of dedicated upstream and downstream wavelength channels. In addition to its operational simplicity, this approach provides dedicated point-to-point optical connectivity to each subscriber with very high bit rate, protocol transparencies, guaranteed QoS, and increased security.

Chapter 3

On the Viability of Implementing A Fully Distributed Control Plane Architecture for EPON

3.1 Introduction

Ethernet-based Passive Optical Network (PON) technology is emerging as a viable choice for the next-generation broadband access network [1-8]. A PON is a point-to-multipoint fiber optical network with no active elements in the signal's path. It consists of a single, shared optical fiber connecting a service provider's central office (head end) to a passive star coupler (SC), which is located near residential customers. Each customer receives a dedicated short optical fiber but shares the long distribution trunk fiber. All transmissions in a PON are performed between an Optical Line Terminal (OLT) and Optical Network Units (ONU's). Traffic from an OLT to an ONU is called 'downstream' (point-to-multipoint), and traffic from an ONU to the OLT is called 'upstream' (multipoint-to-point). Two wavelengths are used: typically 1310 nm (λup) for the upstream transmission and 1550 nm (λd) for the downstream transmission.

In the downstream direction, an EPON operates as a broadcast and select network. The OLT has the entire bandwidth of the channel to broadcast standard formatted 802.3 Ethernet frames to all ONUs. Each ONU extracts those packets that contain the ONUs unique Media Access Control (MAC) address. In the upstream direction, multiple ONUs share the transmission channel. Thus, the ONUs need to employ some arbitration mechanism to avoid collisions.

The IEEE 802.3ah task force is actively standardizing the control and management messages used to control the data exchange between the OLT and the ONUs as well as the processing of these messages through the development of Multi-Point Control Protocol (MPCP). The protocol relies on two Ethernet control messages (GATE and REPORT) in its regular operation. The OLT assigns the Transmission Windows (TWs) via the GATE messages.

In general, the OLT arbitrates the upstream transmissions by allocating an appropriate timeslot/transmission window to each ONU. An ONU is only allowed to transmit during the TW allocated to it by the OLT. Each ONU uses a set of queues to store its Ethernet frames and starts transmitting them as soon as its TW starts. Within each cycle, in order to inform the OLT about its bandwidth requirements, ONUs use REPORT Messages that are also transmitted along with the data in the same TW. The ONU should also account for additional overhead when requesting the next time slot; this includes 8 bytes frame preamble and 12 bytes Inter-Frame Gap (IFG) between two consecutive frames. Between the TW of two ONUs there is a certain guard time “g” needed to account for the laser on and off times, receiver recovery times and other optic related issues. Upon receiving a REPORT, the OLT passes the message to a Dynamic Bandwidth Allocation (DBA) module, which performs the bandwidth allocation computation.

Several bandwidth allocation schemes have recently been reported in the literature ranging for a static allocation to a dynamically adapting scheme based on instantaneous queue size in every ONU [5-8]. To date, the mainstream of these EPON bandwidth allocation schemes as well as the new IEEE 802.3ah EFM Task Force specifications [2] have been centralized–relying on a component in the central office (OLT) to provision upstream traffic. Hence, the OLT is the only device that can arbitrate time-division access to the shared channel. Since the OLT has global knowledge about the state of the entire network, this is a centralized control plane in which the OLT has a centralized intelligence.

Recently, we have proposed a novel Ethernet over Star Coupler-based PON architecture [9]. The architecture uses a fully distributed collision-free DBA scheme in which the OLT is excluded from the implementation of the time slot assignment. It has been shown that in addition to the added flexibility and reliability associated with a distributed control plane, the performance of the proposed decentralized EPON architecture and the associated bandwidth allocation algorithms are as efficient as their centralized counterparts [9].

This paper demonstrates the feasibility of implementing such a novel Ethernet over Star Coupler-based PON architecture that uses a fully distributed time division multiple access arbitration schemes. Specifically, we assess the viability of implementing a distributed control plane architecture in which the OLT is excluded from the implementation of the time slot assignment. In addition to reducing the OLT complexity, by internetworking among connected users (ONUs), the proposed architecture offers efficient shared LAN capability.

To implement a distributed control plane, direct connectivity (communicability) between the ONUs should be in place without imposing any constraint on the PON topology. In the proposed architecture, the ONUs exchange signaling and control information concerning their queue status and their transmission needs amongst themselves. Then, the ONUs simultaneously and independently runs instances of the same DBA algorithm outputting identical bandwidth allocation results. Once the algorithm is run, the ONUs sequentially and orderly transmit their data without any collisions, eliminating the OLT's centralized task of processing requests and generating grants for bandwidth assignment.

3.2. The Proposed Distributed Architecture

Fig. 3.1 shows the general architecture of this approach [9]. As can be seen from Fig. 3.1, a portion of the optical signal power transmitted by an upstream transmitter (λup) toward the OLT will be redirected back and broadcasted to all ONUs. This can be achieved by connecting two ports of a 3xN SC with each other through an optical isolator as shown in Fig. 3.1 [10]. Note that in addition to the conventional transceiver maintained at each

ONU (a λup transmitter and a λd receiver), this approach requires an extra receiver tuned at λup. A baseband direct detection circuit is needed to detect the redirected control channel (λup) in order to recover the control update information.

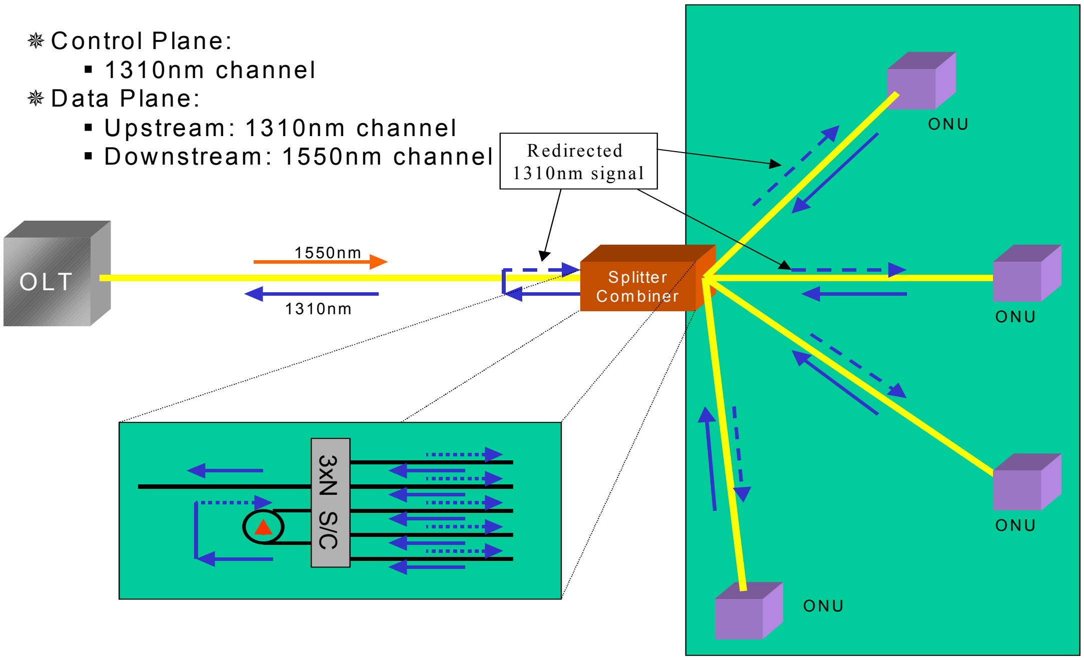


Figure 3.1 Proposed Distributed EPON Architecture

This architecture assumes a cycle-based upstream link, where a cycle is defined as the time that elapses between two executions of the scheduling algorithm. The cycle size can either has fixed, or variable length confined within certain lower and upper bounds to accommodate the dynamic upstream traffic conditions. The cycle is divided into three periods: a static update period (control plane), a fixed waiting period (processing control messages and running the algorithm) and a dynamic transmission period (data plane).

The proposed cycle, along with the details of how the control plane performs the updating process is shown in Fig. 3.1.a in three phases. Each ONU transmits its update control message in its own assigned fixed time slot (first phase). These messages are then combined at the SC and a multiplexed update message is created (second phase). In the third phase, a fraction of the multiplexed control signal is transmitted through the first output port of the SC and propagates to the OLT (which could discard it, make use of it as a synchronization message, and/or process the control information). Another fraction of the multiplexed control signal is redirected back and broadcasted to all ONUs (through the isolator). A baseband direct detection circuit located at each ONU is used to detect the redirected control channel (λupdate). The detected signal is then processed in order to recover the control data information belonging to each of the other (N-1) ONUs. Since there are only two operating communication wavelengths (λup and λd), signaling and upstream transmission take place on the same communication channel (λup) and the periods will appear sequentially as on the top of Fig. 3.2

1. The First Period (Control Plane): The update period is divided into N equal fixed time slots where N is the maximum number of users the network is designed to accommodate. The update period is used for the ONUs to communicate their status and to exchange signaling and control message information with one another. Each ONU uses its own fixed time slot within the update period to transmit its control message. For simplicity, and to avoid collisions, the assignment of these N timeslots follows a fixed Time-Division Multiple Access (TDMA) assignment since control messages are fixed in size.

Note that the control slots in the proposed distributed scheme are all transmitted sequentially in one period (update period). This is in contrast to the centralized schemes reported above [3-8], where the control slot (REPORT Message) of each ONU is transmitted along with the data in the TW allocated to it by the OLT. All control update messages are transmitted as Ethernet frames. Because the signaling information is segregated from the upstream traffic, signaling information can be timelier and complete thus increasing the efficiency of the dynamic bandwidth allocation algorithm. These enhanced DBA algorithms would have the ability to support better QoS characteristics because transmission of the signaling information is not constrained by the shared data/control upstream channel associated with the centralized schemes.

2. The Second Period (Algorithm Execution): The second period of fixed length is a waiting period (no upstream transmissions are allowed during this period) and is used for allowing the ONUs to process the information gathered from the multiplexed control message. Each ONU maintains a table with information about the state of the queues at each other ONU. This information is updated each cycle whenever the ONU receives a new multiplexed control messages from all other ONUs. The DBA algorithm module uses the table maintained at each ONU. Note that instances of same DBA algorithms are executed simultaneously and independently at each ONU.

Note that instances of the same DBA algorithm are executed simultaneously and independently at each ONU. An execution of the algorithm produces a unique and identical set of ONU assignments wi; where wi is the amount of bytes that an ONU is allowed to transmit in its TW during the cycle. It is critical that the algorithm produces a unique outcome for any arbitrary set of inputs. In other words, it should not incorporate any assumptions or randomness to handle exceptions. This is because several instances of it will run locally and independently at each ONU.

3. The Third Period (Data Plane): The third period or (transmission period) is essentially a giant slot used for actual upstream data transmission. During the transmission period, the ONUs follow exactly the allocation scheme the algorithm produced (i.e., their transmissions start at specific times and last for specific bytes) as shown in Fig. 3.2 b. Note that the order of ONUs transmission may be different in each cycle and need not be fixed. It is rather a function of the ONU's traffic demand. This is a major advantage compared to the fixed transmission order proposed in [8].

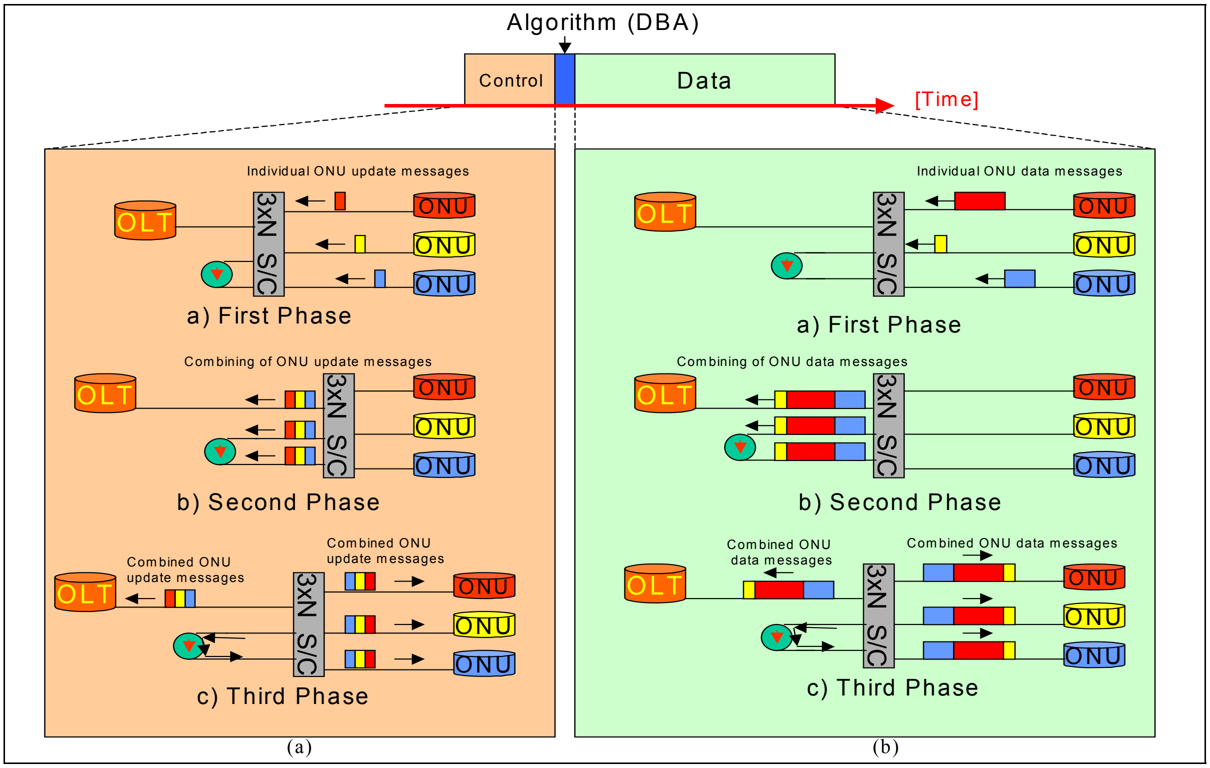


Figure 3.2: (a) Cycle updating process (b) Transmission process.

3.3 Experiments and Discussion

The experimental work presented in this section focuses only on the control plane implementation of the proposed. decentralized EPON architecture. Fig 3.3 depicts the experimental setup of interconnecting four ONUs through a 2x4 bi-directional coupler. Each ONU consists of two sections; a transmitter and a receiver. The upstream transmitter section consists of an HP Data Generator that generates programmable 2^6-1 PRBS NRZ signal at 625 MHz. The NRZ signal directly drives a 1310 nm FP laser (λup) that has a Rise/Fall time of 0.3 ns. Since control messages are fixed in size, the relative timing of the upstream transmitters is programmed to emulate a fixed Time-Division Multiple Access (TDMA) assignment. The laser is followed by an isolator, which protects the laser from the reflected signals. The Receiver section consists of an Agilent photodetector front end and an HP Data Analyzer. At each ONU, both transmitter and receiver sections are coupled by a 3 port optical isolator whose port # 2 is connected to the 4x2 coupler through SMF.

Each ONU transmits its update control message in its own assigned fixed time slot. These messages are then combined at the SC and a multiplexed control signal is created at output port #1 of the 4x2 BDC. Thus, the control signal output power drops down to 1/N of the initial power. This signal is then fed back into the second output port #2 (as input), which is redirected back and broadcasted to all ONUs.

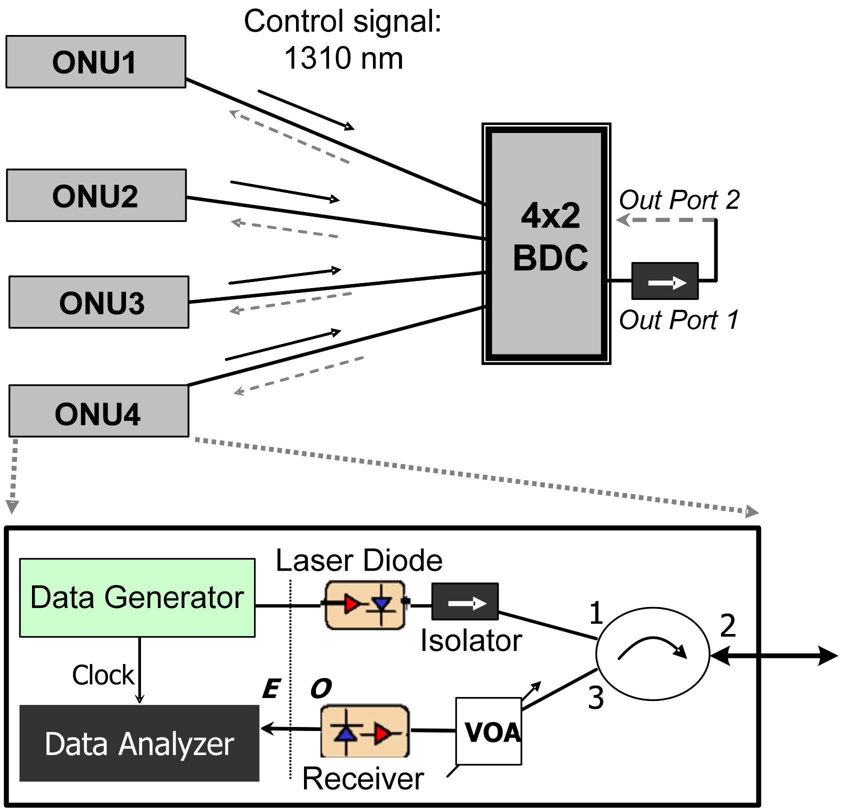


Figure 3.3 Experimental Setup of Decentralized EPON Control Plane

Thus, the power from a given ONU transmitter backs to the corresponding ONU receiver scales as $1/N^2$. Consequently, the main objective of this experiment is to explore whether the SNR of the redirected control signal is still adequate for the ONU receiver to recover the control messages. With the experimental setup shown in Fig.3.3, the receiver sensitivity (minimum power level of the redirected control signal) to achieve a zero BER is measured to be – 38 dBm.

First, we measured the eye diagram of the redirected control signal. As can be seen from Fig.3.4, the eye is wide open with narrow crossover width (low jitter). Thus, the eye establishes a good signal quality indicating that the received control signal experiences little or no distortion. Second, we compare the performance of the proposed decentralized architecture with that of a reference back-to-back configuration. Back-to-back configuration is simply composed of a transmitter with its driver board, followed by a baseband direct detection receiver and an analyzer. For both back-to-back and decentralized configurations, a variable optical attenuator is used to vary the input power to the receiver. The figure of merit (Q factor) can be written in terms of the measured BER values as [11]:

$$BER = \frac{\left[e^{\left(-Q^2/2\right)}\right]}{Q\sqrt{2\pi}}$$

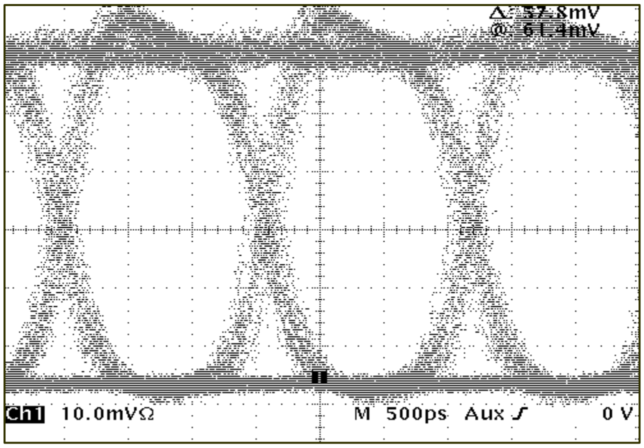


Fig.3.4 Eye Diagram of Received Signal of Decentralized EPON Control Plane

Fig. 3.5 shows received power versus 20 Log Q for both the decentralized and back-to-back configurations. As can be seen from the figure, the performance penalty experienced with decentralized architecture as compared to that of the reference back-to-back configuration is *less than* 0.6 dB. This penalty is mostly due to the simultaneous on/off light sources performances.

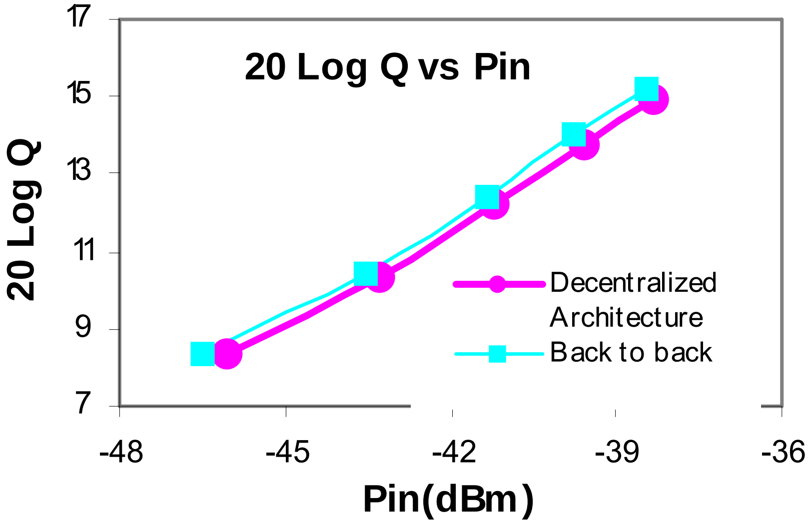


Fig.3.5 20 Log Q vs Received Power

The largest number of ONU’s (N) that can be attached to a single PON is limited by the optical power budget: there has to be sufficient optical power budget to support the huge splitting losses involved in such distributed control plane architecture. With a 1dBm transmitted power/ONU used in our experiment along with –38 dBm measured receiver sensitivity; a power budget of 39 dB is adequate to support 32 ONUs. This number could be further increased with higher transmitter power and receiver sensitivity.

3.4. Conclusion

We have demonstrated the viability of implementing a novel Ethernet over Star Coupler-based PON architecture that uses a fully distributed time division multiple access arbitration schemes. Specifically, we have assessed the viability and performance of a distributed control plane architecture in which the OLT is excluded from the implementation of the time slot assignment. It has been shown that the SNR of the redirected control signal is still adequate for the ONU receiver to recover the control messages. Furthermore, the optical power budget of the proposed decentralized EPON architecture is sufficient to support 32 ONUs.

Chapter 4

A Novel Open Access CWDM-PON Architecture

4.1 Introduction

PONs enhances the fiber penetration towards the end-user side to carry various higher data rate services to residential users. Due to their inherent advantages, e.g., very high bandwidth, protocol transparency, cost factor, coarse wavelength division multiplexed (CWDM) systems are becoming increasingly important in the access networks. By using inexpensive components, e.g., uncooled directly modulated lasers(U-DML), full spectrum CWDM systems presently offers up to 18 wavelengths which are spaced 20-nm between 1270 nm and 1610 nm over low-water peak fiber. Before CWDM, all the works on fiber to the home (FTTH or FTTC, FTTB) stated utilizing of expensive DFB lasers at the OLT. But as it is experimentally demonstrated at [1] where bidirectional CWDM setup make use of 16 U-DMLs operating over low-water-peak fiber (LWPF), inexpensive U-DMLs can replace DFB Lasers in the downstream direction and further reduce the cost of CO Optics.

Government regulations in many countries have mandated a clear demarcation between the network operators who provide physical connectivity and transport data (e.g., telephone and cable companies), and the service providers (SPs) who deliver the content and the services (e.g., Internet SPs). The primary motivation is to have free market competition between SPs, thereby making services cheap and available to end users; and also preventing monopoly by the network operator who owns the right-of-way to lay cable/fiber in a residential area. Such a framework is known as ***open access.***

CWDM-PON utilization on the access networks reveals the fact that residential customers who are on the same CWDM-PON access network can be served with different data link layer technologies. Having only one data link layer technology causes the monopoly of one data link layer at a given access network. Moreover, voice, video and data services may be limited by the capability of chosen data link layer technology**.**
We believe open access shown in [2] is also achieved best with the proposed CWDM-PON architecture since network operators assigns a set of wavelengths to each SPs for their operations. Note that, since each SP operates on different set of wavelengths, SPs' services are transparent to each other

4.2 Proposed CWDM PON Architecture

Proposed architecture is shown in Figure 4.1. In the proposed architecture, up to 8 service providers (SPs) serve 4 neighborhoods over a single CWDM-PON.

In the proposed configuration, each SP has its own data link layer equipment located at the CO and they are assigned a pair of wavelengths (one downstream, one upstream) by the network operator for their operations. Proposed architecture provides a lot of flexibility to SPs. First, since SPs' services are transparent to each other they are free to choose any PON technology (i.e., BPON, EPON, GPON) desired. SPs data rates provided to their end-users both in up- and downstream direction might be independent of each other. In this architecture end-users have service level agreement (SLA) with the SP only.

To multiplex and demultiplex different SPs' wavelengths to the transmission line, a CWDM located at the CO. After multiplexed together, combined downstream wavelengths are split to four copies where each copy is transmitted to different neighborhoods. At each neighborhood, each copy is further split (via 1x8 splitter) to 8 copies before each copy is demultiplexed by a CWDM which is located at the remote node (RN). Location of 1x8 splitter, which could be also a hierarchical set of small size splitters can be determined based on the fiber aggregation points so that the amount of fiber to be deployed is minimal in the access network.

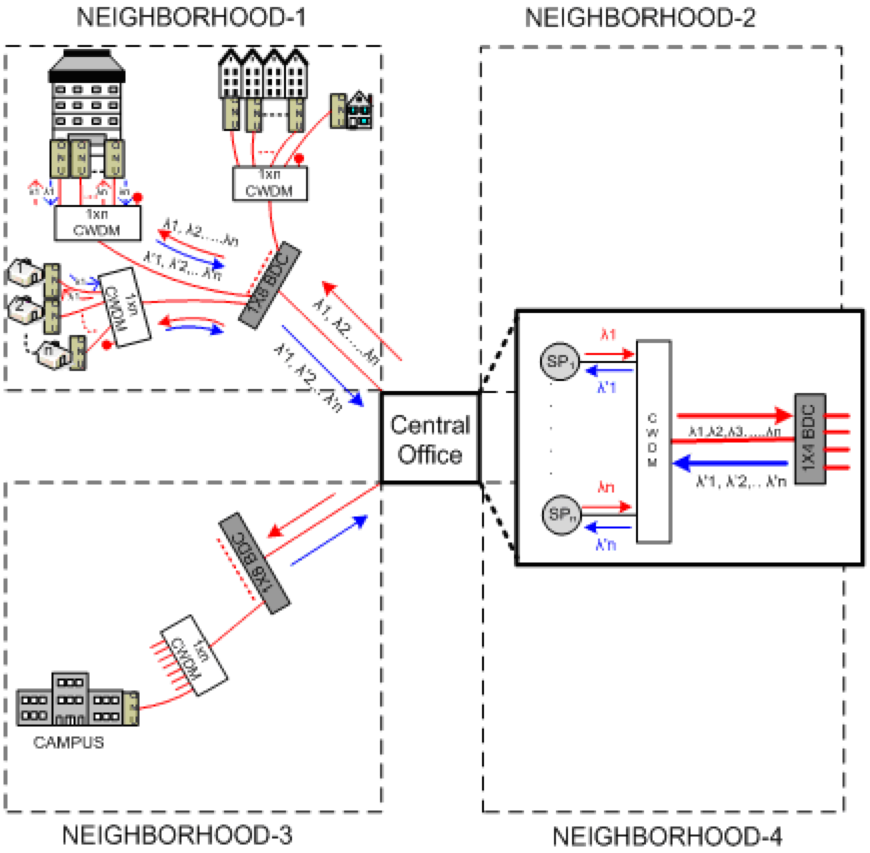


Figure.4.1 CWDM-Based PON Architecture

4.2.1 Power Budget

Power budget of the proposed CWDM PON architecture is the same as 32-split conventional single wavelength PON system except that there is additional loss due to two CWDM equipments used in the proposed architecture. Our calculations using parameters from commercial devices show enough power margins even for 32:1 split. Insertion losses are determined based on the commercial devices [3]. The power losses for all components are listed below: 2*CWDM = 2.4 dB; Fiber loss = 0.25dB/km * 20km

= 5dB; Coupler loss = 10dB (for 8:1), 7dB (for 4:1). Including 2 dB margin, total insertion loss is around 26.4 dB.

4.3 Dynamic Bandwidth Allocation Algorithms in Downstream

In this section, aligned with proposed architecture, performance of two dynamic bandwidth allocation algorithms are compared in downstream over EPON. The fact that well over 95% of all data traffic either originates or terminates as Ethernet traffic that has prompted us to select Ethernet as the potential convergence solution for the proposed access network. Ethernet Passive Optical Network (EPON) is one of the access solutions considered by Ethernet in the First Mile (EFM) Task Force focusing on direct support of Ethernet services.

All users are assumed to have same service level agreement with an SP. Thus, to achieve fair queuing across different end-users, a fairness algorithm needs to be implemented at the OLT. A number of algorithms have been proposed to achieve fair queuing across different independent network flows in the literature, e.g., Deficit Round Robin (DRR) [4], Weighted Fair Queuing (WFQ) [5], and Fair Queuing with Round Robin [6]. Considering the fact that DRR is an important fair queuing algorithm that is used in CISCO's Gigabit Switch Router 12016, we are going to implement DRR for fair queuing algorithm at the OLT.

4.4 Deficit Round Robin Algorithm (DRR)

DRR algorithm works in rounds. A round is one complete cycle through all backlogged buffers of sending assigned amount of bytes per buffer. Note that the time duration of a round (cycle) depends on the number of backlogged buffers. Each buffer is allowed to send out packets in the first round no more than *Quantum*. *Quantum* is a fixed amount of bits. If there are no-more packets in queue "*i*" after the queue has been serviced a state variable called $Deficit_i$ is reset to 0. Otherwise the remaining amount is stored in the state variable $Deficit_i$. In subsequent rounds, the amount of bandwidth usable by this queue is the sum of $Deficit_i$ of the previous round and *Quantum*. To avoid examining empty queues, an auxiliary list called *ActiveList* which is a list of indices of queues that contain at least one packet is used. For detailed explanation of DRR readers are referred to the [4].

ITU G.114 recommends that the maximum tolerable delay of voice traffic in an access system be 1.5 ms. When implementing DRR, since the low packet processing overhead of DRR (O (1)) is achieved when *Quantum* is bigger than or equal to maximum packet size, P_{max}[4], the lower limit of *Quantum,* Q_{min}, is set to P_{max}. Upper limit of Quantum, Q_{max}, is bounded with maximum tolerable delay of voice traffic.

In this work, maximum buffer slot, S_{MAX}, is defined as maximum transferable window size for a given user in one round. S_{MAX} limit is given as follows;

$$S_{MAX} = Q_{max} + P_{max} - 1, \quad i \in 1 \ldots N$$

P_{max} -1 is the maximum deficit that can be left from previous round (cycle) for any buffer. Maximum round slot (S_{Round}), which is one iteration of processing each buffers that utilizes its assigned slot maximally, can be found with the following formula;

$$S_{Round} = N * S_{MAX}$$

where N is number of downstream buffers at OLT. After setting the upper and lower limit of *Quantum*, we decided to use Quantum value of 2389 bytes for DRR algorithm so that the maximum scheduling delay of a packet at a buffer is equal to ~500 microseconds as in [2].

4.5 Priority queuing embedded DRR at CWDM PON

Using DRR at the OLT provides only fairness but not differential treatment to different class of service (CoS) traffics. Our goal at this part is to integrate a QoS scheduler to DRR module, so that SPs not only provides fairness among different subscribers' traffic, but also provides differential treatment to different CoS traffics within a subscribers' traffic. In this work, all subscribers are assumed to have same SLA with the SPs.

As for QoS algorithm, where the performance of Strict Priority Queuing (SPQ), and Fair Non-Strict Priority Queuing (F-NSPQ) is compared when they are both embedded to DRR.

The scheduler block diagram is shown in fig.4.2. When a packet enters into the OLT it will be first classified based on the destination MAC address. And then it will be further classified based on the CoS and put into corresponding queues of the buffer.

In one complete cycle (round), DRR scheduler assigns a transmission window called ***Quantum**,* to each active buffer for its downstream transmission. If there is *deficit* from the previous cycle, then deficit will be added to ***Quantum*** and resultant window will be called ***BufferWindow.*** ***BufferWindow*** is shared among 3 sub-queues (Q_0, Q_1 and Q_2) based on QoS algorithm implemented.

4.6 DRR integration with Strict Priority Queuing (SPQ)

Strict Priority Queuing (SPQ) is used as a default scheduling algorithm in IEEE 802.1D-compliant bridges and switches and is defined in P802.1D. In SPQ, arriving **v**oice packets displaces data packets from the tail of data queue if buffer is full without any condition. In case buffer is full while there exist no data packets in the buffer then arriving voice packets displace video packets, if any. Voice packets will be dropped only if buffer is full of voice packets and there is no place for newly arriving voice packets. When scheduling an active buffer which can deliver ***BufferWindow*** amount of packets, scheduler first delivers all the voice packets. Then the remaining window from ***BufferWindow***, if any, is going to be used for scheduling of all video packets and if there is still some window left from ***BufferWindow,*** it will be used for data sub-queue traffic. After servicing a given buffer, if there is still some window left to be used, but all the buffer packets are scheduled, then deficit will be reset. But if there is still a packet waiting at one of sub-queues which's size more than the window left then that window will be used at the next round as deficit. The biggest problem with SPQ algorithm that data packets can wait forever at the data sub-queue at the high load just because ***BufferWindow*** might be used with all voice and video packets and nothing is left for the data traffic.

4.7 Non-Strict Priority Scheduling (NSPQ)

This algorithm is similar to the one used in [9]. Algorithm starts transmitting all voice packets first then video packets and then data packets. After servicing a buffer, algorithm keeps record of the total number of packets that are waiting at each sub-queue of the buffer. After scheduler completes its round and come back to schedule same buffer again, it checks its record to find out the packets from the each sub-queue which are waiting from the previous round since they will be transmitted next and subtract total value of all these packets from the ***BufferWindow.*** After subtracting it, if there is still some window left from the ***BufferWindow*** that window is going to be used for the packets which came between first and second round scheduling time of the same buffer. Assume a buffer has in the record $W1_{vo}$ amount of voice, $W1_{vi}$ amount of video and $W1_{dt}$ amount of data and W2vo, $W2_{vi}$ and $W2_{dt}$ amount of packets came between two scheduling time of the buffer. Assume scheduler has enough ***BufferWindow*** to schedule all these packets. The order will be $W1_{vo}+ W2_{vo}$, $W1_{vi}+ W2_{vi}$ and $W1_{dt}+ W2_{dt}$. So every time voice has the highest priority while data is the lowest. But if there is simply not enough windows to transmit $W2_{vo}$, $W2_{vi}$ and $W2_{dt}$, then only $W1_{vo}$, $W1_{vi}$ and $W1_{dt}$will be transmitted.

Both QoS algorithms introduced above use same OLT Buffer Block Diagram shown in Figure 4.2.

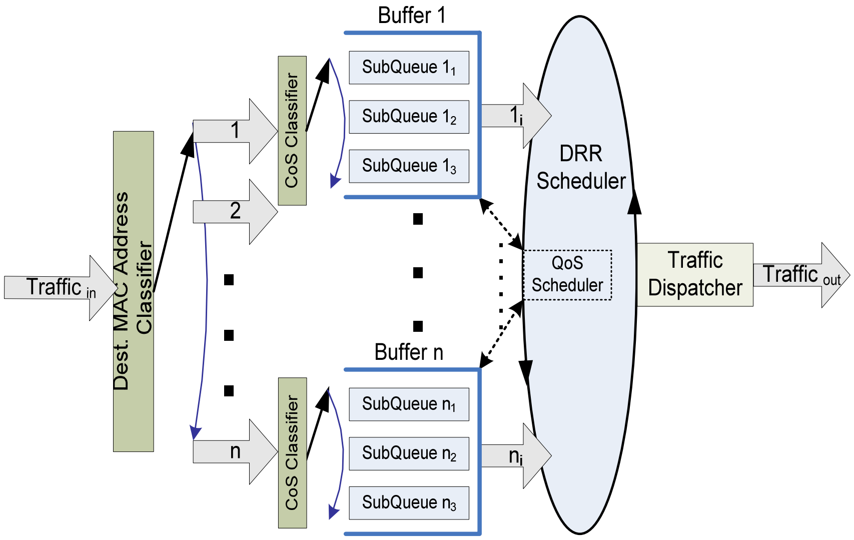


Figure 4.2 OLT Buffer Block Diagram

4.8 Performance Analysis

In this work, OLT is connected to 16 ONUs and corresponding to each ONU houses 16 downstream buffers with 1MByte size each. Each buffer is composed of 3 different sub-queues to realize QoS for different CoS traffic. Those queues are called as Q_0 (CBR), Q_1 (VBR) and Q_2 (Best Effort data, (BE)).

The traffic used for Q_1 and Q_2, is self similar traffic with long range dependence with variable packet size ranging from 64 byte to 1518 bytes, while for Q_0, T1 connection is emulated which has the constant rate of 4.48 Mbps as in [7]**.** In our traffic model, downstream traffic (load) is generated as follows; for 0.1 load (100 Mbps), 100 Mbps is divided equally between 16 downstream buffers, ~6Mbps. Since at every load each downstream buffer receives constantly 4.48 Mbps of Q_0 traffic, the remaining traffic (~6-4.48Mbps) is distributed equally between Q_1 and Q_2 traffic. Thus, at higher loads load increase is reflected on $Q_1$and Q_2 traffic where each gets equal share while Q_0 traffic stays constant (4.48Mbps) at every load.

Description	Value
Number of ONUs	16
EPON Line Rate	1 Gbits/s
Distance between OLT and ONUs	20km
Quantum	2389 Byte
Number of priority classes at each buffer	3
Number of buffers at the OLT	16
Buffer Size at OLT (For each subscriber)	1 MByte

Table 4.1 Simulation Parameters

4.9 Simulation Results

Figure 4.3 shows the average buffer delay for voice traffic at SPQ and NSPQ algorithms. Even under heavy load conditions figure 4.3 proves that average voice delay is below the maximum tolerable delay requirement (1.5 ms) of voice packets at the access networks for both algorithms.

Since voice packets which come between 2 consecutive rounds might wait next round to be transmitted at NSPQ, voice delay is the higher at NSPQ at heavy loads.Figure 4.4 shows average buffer delay of video traffic for SPQ and NSPQ algorithms. Delay is almost same for both algorithms under different loads.

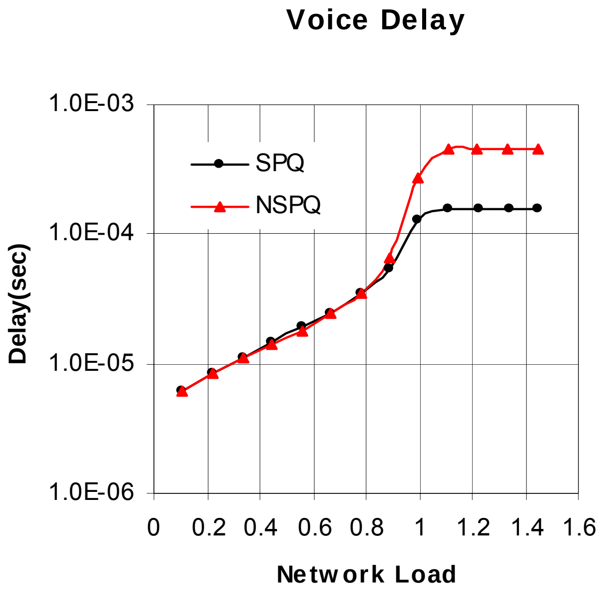


Figure 4.3 Mean Buffer Delay of Voice packets at SPQ and NSPQ versus ONL

Figure 4.5 shows average buffer delay of data traffic for SPQ and NSPQ algorithms. Average data delay at NSPQ is slightly smaller than at SPQ algorithm between the loads of 0.5 and 1. This is due to, in a given round, all scheduled data packets are sent out in NSPQ. As for SPQ, scheduled data packets waits as long as there some unscheduled vide or voice packets present in the buffer.

Figure 4.5 Mean Buffer Delay of Data packets of SPQ and NSPQ vs. ONL

Figure 4.6 shows data drop ratio at SPQ and NSPQ algorithms as a function of ONL. As expected, when the load is less than 1, drop ratio for data packets less in NSPQ than SPQ. In SPQ, bursty fluctuation of incoming traffic might fill up downstream buffer very quickly, then newly arriving voice or video packets drops data packets from the tail of the queue. This causes more data packets drop in SPQ as compared to NSPQ.

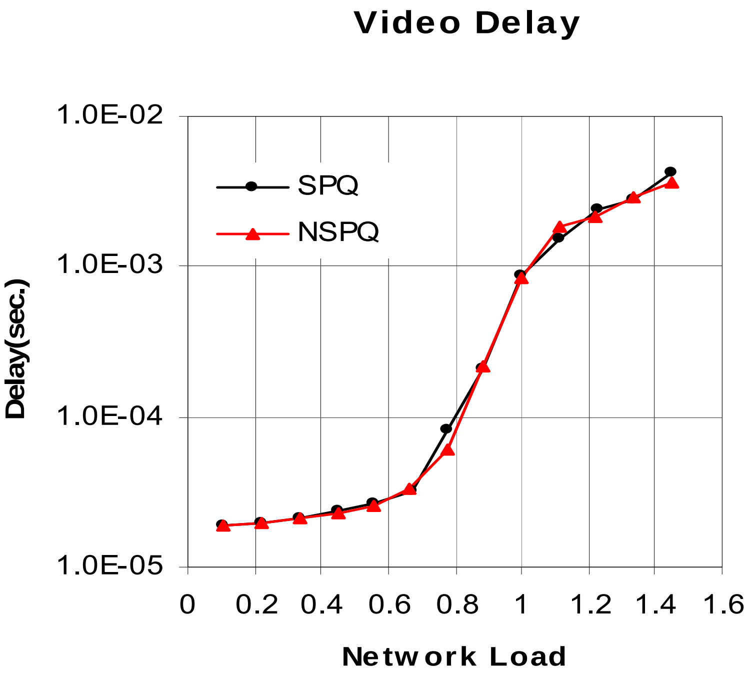


Figure 4.4 Mean Buffer Delay of Video packets of SPQ and NSPQ vs. ONL

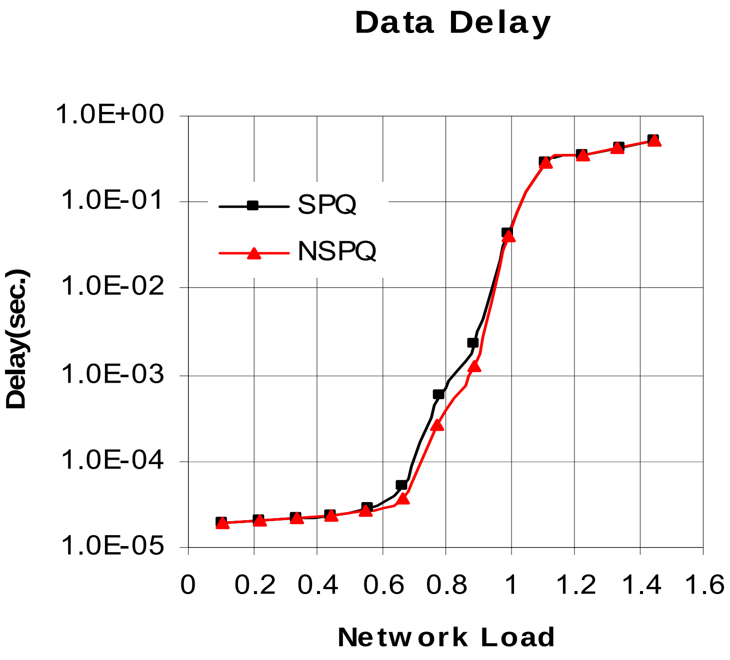


Figure 4.5 Mean Buffer Delay of Data packets of SPQ and NSPQ vs. ONL

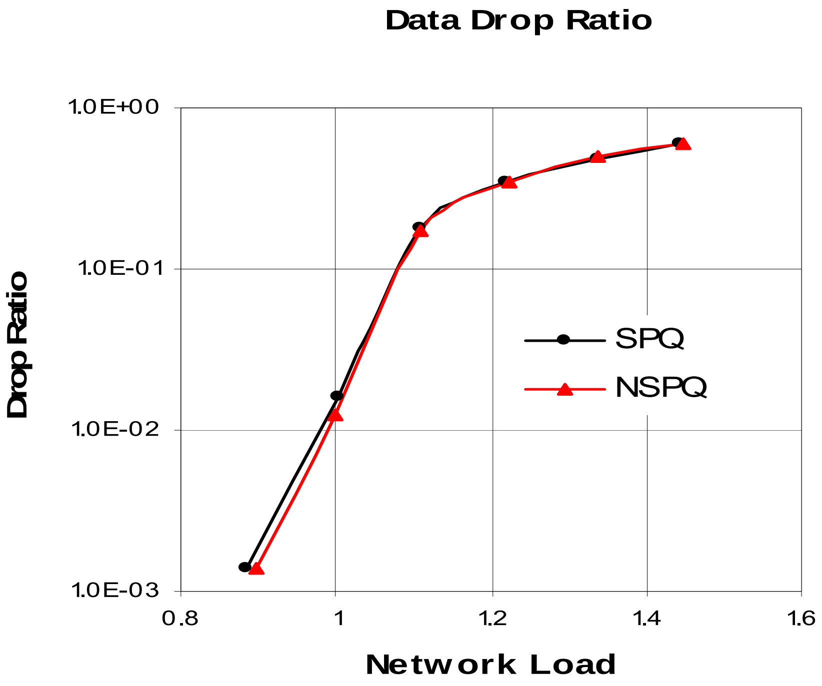


Figure 4.6 Data Drop Ratio at SPQ and NSPQ algorithms vs. ONL

4.10 Conclusion

This work has proposed a cost effective CWDM-PON architecture that is open to every service provider. By using CWDM-PON on the access networks, residential customers can be served with same/different data link layer technologies with the same/different SPs with same/different data rates based on SLA between SP's and end-users. In proposed CWDM-PON architecture, since every SP's services are transparent to each other, each SP is free to select among different PON technologies (BPON, EPON, and GPON). Supported by the CWDM-PON scheme, we have integrated DRR to two different QoS algorithms (SPQ and NSPQ) to not only provide fairness among the end-users but only differential treatment to different CoS flows.

Chapter 5

A Simple and Cost Effective Ring-Based Local Access C/DWDM-PON Architecture for Supporting a Truly Shared LAN Capability

---

5.1 Introduction

Passive Optical Network (PON) technology is emerging as a viable solution for next-generation broadband access networks [1-5]. A PON connects a group of Optical Network Units (ONUs) located at the subscriber premises to an Optical Line Terminal (OLT) located at the service provider's central office (CO). Among the various PON-based FTTH schemes, single channel Time-Division Multiplexed PON (TDM-PON) and multi-channel Wavelength-Division Multiplexed PON (WDM-PON) architectures are the two most viable candidates.

TDM-PON supports a single wavelength channel in the downstream direction (OLT to ONUs) and another wavelength in the upstream direction (ONUs to OLT). Current TDM-PONs including broadband PON (BPON), gigabit PON (GPON), and Ethernet PON (EPON) provide higher bandwidth than traditional copper-based access network solutions and seem to offer a satisfactory solution for present bandwidth demands. However, since all users attached to a TDM-PON share a single transmission channel, the average dedicated bandwidth assigned to each user in either direction is usually limited to a few percent of the channel capacity, i.e., a few tens of Mb/s. As the number of end users attached to a given access network grows each year, and as their usage patterns evolve to include more bandwidth-intensive content and peer-to-peer applications, there emerges an acute need for broadband access solutions that can upgrade current TDM-PONs and provide each user with at least a 100 Mb/s, symmetric, and guaranteed bandwidth [5].

WDM-PONs are emerging as the most promising future access solutions that can provide evolutionary upgrade to existing TDM-PONs. These schemes can support multiple wavelengths in either or both the upstream and downstream directions. Traditional WDM-PON systems allocate a separate pair of dedicated upstream and downstream wavelength channels to each subscriber, enabling the delivery of 100 Mb/s or more of dedicated bandwidth per subscriber or ONU. In contrast to traditional TDM-PON architectures, WDM-PON has a unique ability to support any future bandwidth upgrade without altering the physical infrastructure. In addition to its operational simplicity, WDM-PONs provide dedicated point-to-point optical connectivity to each subscriber with bit rate and protocol transparencies, guaranteed QoS, and increased security. Moreover, WDM-PONs can support the transmission of any service or service mix, including Gigabit Ethernet, ESCON, FICON, Fiber Channel, OC-N, ATM, and FDDI, at any bit rate.

Despite these aforementioned numerous crucial advantages, traditional WDM-PON architectures suffer from several limitations including inability to efficiently utilize network resources and to cope with the dynamic and bursty traffic patterns of the emerging integrated triple pay services [5-6]. In addition, since most of TDM/WDM PON architectures are traditionally deployed as tree topologies, end users attached to these PONs cannot directly communicate with each other. The former limitation is exacerbated when some wavelength channels are heavily loaded while others are underutilized or are totally idle. In this case, unused dedicated channel capacities of those lightly loaded/idle subscribers cannot be shared by any of the other heavily loaded users attached to the PON, leading to the waste of scarce network resources. Therefore, to increase the total throughput, future WDM-PON architectures must support dynamic bandwidth allocation (DBA) and sharing.

To address the former problem, several WDM-PON architectures and protocols that dynamically manage and allocate bandwidth in both time and wavelength dimensions have been proposed recently [6-11]. Most of these schemes, however, are costly and assume complex OLT and ONUs setups, which require tunable, or array of fixed transceivers, or both, WDM filters, and wavelength-band-selective receivers. Furthermore, schemes that support dynamic wavelength sharing [6-7], where additional wavelength channels are added to accommodate the fraction of bursty downstream traffic that may exceed the user’s dedicated downstream wavelength channel rate, are still falling short of addressing the fundamental problem of the inefficient utilization of network resources. This is because the unused capacities of those lightly loaded/idle dedicated downstream wavelength channels are still being wasted.

The latter problem has recently received considerable attention due to the rising importance of supporting virtual private connections among end users (for instance, branch sites in a business enterprise). In this regard, several physical layer LAN emulation schemes have been proposed to achieve intercommunication among the ONUs, but only within the context of TDM-PONs [12-14]. Most recently, fewer schemes have been proposed to achieve intercommunication among the ONUs within a WDM-PON infrastructure [15-16]. In addition to the added cost and complexity, most of these schemes, however, suffer from poor scalability due to high splitting losses as the redirected LAN signals traverse through the star coupler/AWG once or twice, resulting in lower power budget that limits the number of ONUs that can be attached to a single PON. In general, achieving intercommunication among subscribers within a tree-based WDM-PON setup is a lengthy and complex process that requires much more resources than those needed for a TDM-PON.

The purpose of this paper is to propose a simple and cost effective local access WDM-PON architecture that addresses some of the limitations of conventional tree-based WDM-PON architectures including supporting dynamic allocation of network resources as well as a truly shared LAN capability among end users. Specifically, this work proposes and devises a novel ring-based local access WDM-PON architecture that efficiently supports dynamic allocation of wavelengths/timeslots and sharing traffic as well as truly direct private connections among PON end users. The proposed architecture combines the salient features of both traditional static WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and increased security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end users).

Unlike a typical WDM metro-access ring network, where the feeder fiber of a PON is replaced with a metro fiber ring that interconnects the hub and access nodes, the proposed architecture interconnects WDM ONUs via a short distribution fiber ring in the local loop but allows them to share the feeder fiber for long reach connectivity to the OLT. This architecture is well suited for an autonomous access environment such as a private financial enterprise where supporting secure, reliable, and direct intercommunications among end users is required and where several high-end users are closely dispersed within a 1-4 km diameter area.

Specifically, we develop an OLT-based shared wavelength/timeslots assignment and scheduling (SWS) algorithm along with a fully distributed LAN DBA scheme that dynamically allocate network resources among end users fairly and efficiently. To achieve this objective, overall network resources, including downstream, upstream, and LAN wavelength channels, are load-balanced and efficiently utilized via traffic-engineered routing of user's traffic. The algorithm is QoS-aware via assigning dedicated wavelength channels for higher priority traffic (e. g., voice and video) and limiting dynamic wavelength assignment and sharing to best-efforts traffic only.

5.2 Proposed Ring Based C/DWDM PON Architecture

Figure 5.1 illustrates the proposed ring-based WDM-PON architecture. An OLT is connected to N WDM ONUs (this work assumes N = 16) via a 20-km trunk feeder fiber, a passive 3 port optical circulator, and a small fiber ring. To cover the same local access area as in the tree-based architecture [1-5], the small ring at the end of the trunk is assumed to have a 1-4 km diameter. The ONUs are joined by point-to-point links in a closed loop around the access ring. The links are unidirectional: both downstream and upstream signals (combined signal) are transmitted in one direction only. Each ONU is assigned a single dedicated wavelength for both downstream and upstream transmissions. Direct intercommunication among ONUs is achieved via an additional local control/LAN wavelength channel, λ_{LAN}, which is terminated, regenerated, and retransmitted at each ONU.

Downstream/upstream wavelengths as well as the LAN wavelength are spaced 200-GHz apart in the 1530-1565 nm standard C-band. Since the total number of required wavelength channels in the proposed architecture is assumed to be 17 (corresponding to 16 downstream/upstream wavelength channels plus an additional local control/LAN wavelength channel), these channels can also be allocated over the 1270-1610 nm CWDM spectrum that can offer up to 18 available channels with 20-nm spacing, as defined in ITU TG.694.2. Thus, the overall system cost can be reduced via the utilization of low cost commercially available CWDM components. To scale beyond the 16 ONUs, the number of downstream/upstream wavelength channels can be doubled or quadrupled by reducing the channel spacing in the C-band to 100 or 50-GHz, respectively, Provided that the overall system power budget is still satisfactory.

The OLT houses an array of N fixed transmitters (Tx) and another array of N+1 fixed receivers (Rx), passive 3 port optical circulator, and a commercially available low cost thin film-based DWDM multiplexer/demultiplexer with channel dependent insertion losses between 0.3 and 5 dB. Each Tx/Rx pair corresponds to one ONU and utilizes the same wavelength for transmitting and receiving downstream and upstream traffic, respectively. The extra receiver (N+1) located at the OLT is used to detect the local control/LAN channel. Each ONU has a Tx/Rx pair which is matched to the corresponding pair at the OLT and another Tx/Rx pair for transmitting and receiving the local LAN channel, λ_{LAN}. In addition, each ONU houses a commercially available low cost four-port thin film filters-based fixed optical add-drop multiplexer (OADM), where two wavelengths (corresponding downstream/upstream and LAN wavelengths) are dropped and added at each node.

The DWDM downstream signal is coupled to the ring via port 3 of the optical circulator. After recombining it with the re-circulated LAN signal via a 2x1 WDM combiner (placed on the ring directly after the optical circulator), the combined signal then circulates around the ring (ONU_1 through ONU_N) in a drop/add and go-through fashion. For instance, at the first ONU, the dedicated downstream wavelength channel λ_1 along with the re-circulated control/LAN channel are dropped and processed; then, the dedicated upstream signal λ_1 along with the regenerated control/LAN channel are added. Finally, at the last node (ONU_N), wavelengths λ_N and λ_{LAN} are dropped/added. Thus, the DWDM downstream signal is terminated at the last node.

The combined DWDM upstream and LAN signals emerging from the last ONU at the end of the ring is split into two components via a (10:90) 1x2 passive splitter placed on the ring directly after the last ONU. The first component (90 percent) is directed towards the OLT via circulator ports 1 and 2, while the second component (10 percent) passes first through a band rejection filter that terminates the DWDM upstream signal. The second component emerging from the band rejection filter, the LAN signal, is allowed to re-circulate around the ring after recombining with the downstream signal (originating from the OLT) via the 2x1 WDM combiner (multiplexer). The first component of the combined DWDM upstream and LAN signal is received and processed by the array of N+ 1 fixed optical receiver (housed at the OLT). Specifically, each of the N upstream optical receiver detects the corresponding upstream signal and recover the MAN/WAN traffic, while the LAN optical receiver, as will be explained below, process the control messages and may discard or process the LAN traffic provided that, as will be shown below, it carries upstream traffic as well.

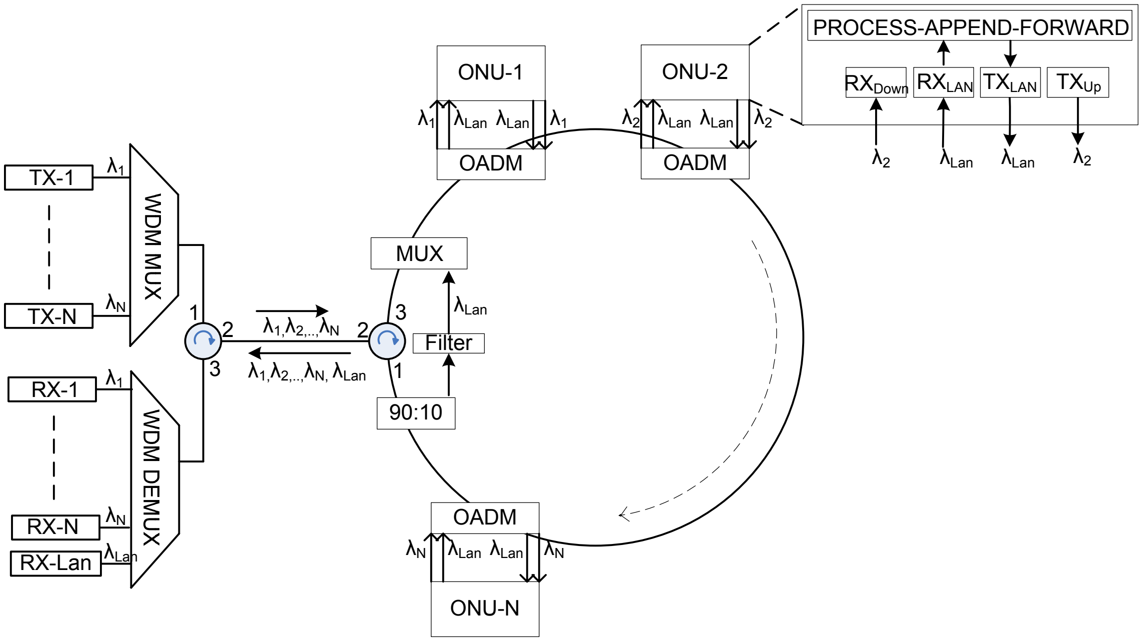


Figure 5.1 Proposed Ring Based C/DWDM PON Architecture

5.3 Distributed LAN DBA Scheme

5.3.1 Overview of Centralized DBA Schemes

In TDM PON, intercommunication among ONUs is provided through OLT. Thus, when ONU_i communicates with ONU_j, ONU_i sends its traffic to OLT via upstream channel, and then OLT reroutes and broadcast that traffic back to all via downstream channel. Thus upstream DBA plays a pivotal role for inter ONU (LAN) communication

Several centralized tree-based upstream DBA schemes have recently been reported in the literature [9-13]. An OLT-based polling scheme, called *Interleaved Polling with Adaptive Cycle Time* (IPACT) based on *Grant* and *Request* messages, has been presented in [9].

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

$$B_{Granted} = \begin{cases} R_i & If R_i \leq B_{MAX} \\ B_{MAX} & If R_i > B_{MAX} \end{cases} \tag{5.1}$$

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

$$B_{max} = \frac{1}{N}[R_{EPON}(T_{MAX} - (N*T_G)] \tag{5.2}$$

where N is the number of ONUs, T_G is the guard band slot, and R_{EPON} is EPON line rate.

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

Note that both of these limitations directly effect performance of not only upstream traffic but inter-ONU traffic also since inter ONU traffic are carried via upstream channel before rerouted and broadcasted to all ONUs by OLT.

5.3.2 Proposed LAN-DBA Scheme Under a Distributed Control Plane

The proposed scheme assumes a cycle-based LAN link, where the cycle size can be either fixed, or variable length confined within certain lower and upper bounds to accommodate the dynamic LAN traffic conditions. During a given cycle, each ONU transmits its control (REPORT) message (within its assigned time slot) around the ring from one node to the next, where it is finally removed by the source ONU after making one trip around the ring. It typically contains the desired size of the next timeslot based on the current ONU’s buffer occupancy. Since these message frames are typical Ethernet frames, the ONU should also account for additional Ethernet overhead when requesting the next time slot; this includes 8-byte frame preamble and 12-byte *Inter Frame Gap* (IFG) associated with each frame. Note that since this overhead is typical to Ethernet frame, there is no additional overhead due to this architecture. Thus, this framing has no impact on our network beside what is standard to any EPON.

Since the REPORT messages are processed and retransmitted at each node, ONUs can directly communicate their status and exchange signaling and control message information with one another.

Each ONU maintains a database that contains the state of the queues of all the ONUs. This information is updated each cycle whenever the ONU receives new REPORT messages from all other ONUs. The DBA module housed at each ONU uses this information to calculate a new set of time slot assignments at each cycle. The ONUs sequentially and independently run instances of the same LAN DBA algorithm outputting identical bandwidth allocation results. The execution of the algorithm at each ONU starts immediately once all REPORT messages have been collected. Thus, all ONUs must execute the LAN DBA algorithm prior to the expiration of the current cycle so that bandwidth allocations scheduled for the next cycle are guaranteed to be ready by the end of current cycle. Note that normally a cycle starts after the LAN DBA calculation for ONU timeslots is completed. Each cycle consists of N timeslots, where N is the number of ONUs in the network. Once all ONUs of a cycle completed their transmissions, then that cycle is over. An execution of the DBA algorithm produces a ***unique*** and identical set of ONU assignments. It is critical that the algorithm produces a unique outcome for any arbitrary set of inputs. Once the algorithm is executed, the ONUs sequentially and orderly transmit their data without any collisions, eliminating the OLT's centralized task of processing requests and generating grants for bandwidth allocations.

It is important to emphasize that maintaining proper time synchronization between ONUs is required for the appropriate operation of the distributed DBA algorithm. In general, this is always the case, as all ONUs are synchronized to a common reference clock extracted from the OLT downstream traffic. Clocking information, in the form of a synchronization marker, is included at the beginning of each downstream frame cycle. The synchronization marker is a one-byte code that is transmitted every 2 ms to synchronize the ONUs with the OLT [17]. The TDM controller at each ONU, in conjunction with timing information from the OLT, controls LAN transmission of the variable-length packets within the dedicated time slots.

Note that in this work, as will be shown below, LAN data might be a combination of native local LAN traffic, transient downstream (TDS) traffic, and transient upstream (TUS) traffic. TDS traffic is defined here as downstream traffic (the fraction of bursty downstream traffic that may exceed the user's dedicated downstream wavelength channel rate) destined to a given access node ONU_i but terminated at a different transient access node ONU_j. This is because TDS traffic is always transported via a non-dedicated downstream wavelength channel λ_j other than its own dedicated downstream wavelength channel, λ_i, which is typically pre-assigned to transport native ONU_i's downstream traffic. Once the TDS traffic is terminated at ONU_j, it is handled as transient LAN traffic whose new source is ONU_j and final destination is ONU_i. The TDS traffic along with

ONU_j's local LAN traffic are first buffered at the corresponding ONU_j's LAN queue, and are then retransmitted over the LAN/control wavelength channel around the ring to their final destinations, within the same granted ONU_j's LAN timeslot.

Similarly, TUS traffic is defined here as upstream traffic (the fraction of bursty upstream traffic that may exceed the user's dedicated upstream wavelength channel rate) originating from a source node ONU_i and destined to OLT, but transported via the LAN/control channel, λ_{LAN}. TUS traffic can be removed from the LAN channel either by source ONU or by ONU1. Due to reason explained below, if removed by the first ONU, since it will lead to better LAN channel utilization, this work will be done by the ONU1.

Since LAN transmission is based on a TDMA scheme, inter-ONU traffic including local LAN data and control messages as well as TDS and TUS traffic, are all transmitted within the same granted timeslot. Thus, devising an efficient and, primarily, a fully distributed LAN DBA algorithm, where the distant OLT is excluded from the process of assigning ONU's timeslots, is central to the successful implementation of dynamic wavelength sharing for both upstream and downstream traffic. Note that at the initial stage of deploying the WDM-PON access network, the line rate of the LAN wavelength channel can be set to either 1Gb/s, 2.5 Gb/s, or 10Gb/s, according to the predictions of worst case scenario of network traffic demands.

To carry out a general and fair comparison between the performances of a centralized DBA scheme and the proposed distributed scheme, this work uses the centralized limited service scheme reported in [9], along with the appropriate changes needed to accommodate the distributed architecture, as the basis for the decentralized DBA scheme presented here. As mentioned above, to globally optimize the bandwidth allocation process, the proposed DBA algorithm execution is performed only after each ONU receives and processes all other ONUs requests.

Based on bandwidth demands, ONUs can be classified into two groups, namely: lightly loaded ONUs that have bandwidth demands less than B_{MAX}; and heavily loaded ONUs that have bandwidth demands more than B_{MAX}. During each cycle, the DBA module must keep track of the unclaimed bandwidth from the set of lightly loaded ONUs. It then must redistribute this excess bandwidth to other heavily loaded ONUs based on their requested bandwidth, i.e. two ONUs requesting bandwidths B1 and B2 more than B_{MAX} will be assigned excess bandwidths proportional to B1 and B2.

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

$$B_{Cycle\_Remainder} = \sum_{i}^{L} (B_{MAX} - R_i) \qquad (5.3)$$

In equation 5.3, L stands for the number of lightly loaded ONUs. The heavily loaded ONUs with $R_i > B_{MAX}$ will require a total over the limit cycle bandwidth:

$$B_{Cycle\_OverLimit} = \sum_{i}^{H} (R_i - B_{MAX}) \qquad (5.4)$$

In equation 5.4, H stands for the number of heavily loaded ONUs. The total remainder cycle bandwidth can be fairly distributed amongst the heavily loaded ONUs to expand their maximum transmission window as follows [10]:

$$\Delta B_i^{extra} = \mathrm{B}_{Cycle\_Remainder} \left[\frac{\mathrm{R_i} - B_{MAX}}{B_{Cycle\_OverLimit}} \right] \tag{5.5}$$

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

$$B_{GH} = \Delta B_i^{extra} + B_{MAX} \tag{5.6}$$

If R_i is the requested bandwidth of $\mathrm{ONU_i}$, $B_{Granted}$ is the bandwidth granted using the proposed limited service-based distributed DBA scheme (Eqs. 3.3 and 3.4), then $B_{Granted}$ can be expressed as:

$$B_{Granted} = \begin{cases} R_i & \text{If } R_i \le B_{MAX} \\ R_i & \text{If } R_i > B_{MAX} \ \& \ B_{Cycle\_Remainder} \ge B_{Cycle\_OverLimit} \\ B_{GH} & \text{If } R_i > B_{MAX} \ \& \ B_{Cycle\_Remainder} < B_{Cycle\_OverLimit} \end{cases} \tag{5.7}$$

Note that the lightly loaded ONUs ($R_i < \mathrm{B}_{MAX}$) can be scheduled instantaneously “on-the-fly” without waiting for DBA module to perform its end of cycle computations. Whereas, the heavily loaded ONUs ($R_i > \mathrm{B}_{MAX}$) will have to wait until all REPORT messages have been received and the DBA algorithm has computed their bandwidth allocations. Thus, lightly loaded ONUs can be scheduled ahead of heavily loaded ones. The drawback of this approach is that it introduces some idle time between two consecutive cycles, n and n+1, provided that the ONUs follow the conventional approach [9-13] and transmit their REPORT messages after data frames. This idle time, during which the PON channel is not utilized, is equal to the time of one trip around the ring (to collect the REPORT) plus the time to perform the DBA computation.

To eliminate this idle time (as proposed above), for a given time slot, the REPORT message is always transmitted before data frames. Assuming that the duration of the time slot allocated only to the **TUS** transmission of the last ONU (ONU that is scheduled to transmit last) is always longer than idle time. In this case, all other ONUs would have an ample time to complete the execution of the DBA algorithm prior to the completion of the last ONU's data transmission (within the same cycle). This ensures that the process of bandwidth allocations to cycle n+1 is always executed before the expiration of cycle n.

The idle time can not be avoided if the last ONU do not have any TUS traffic to transmit and it is equal to one round trip time around the ring (~15μs). When compared to the idle time at limited scheme reported above (80μs idle time due to accumulation of guard times for 16 ONU system), this idle time is still fairly lower.

Note that, in contrast to centralized architectures where the order of the ONUs transmission is fixed (sequential) in each cycle, the distributed architecture has the added flexibility of varying the order of the ONUs transmission according to ONUs traffic demands and priority. Thus, the order of ONUs transmission may be different in each cycle and need not be fixed. As mentioned above, this work assumes that lightly loaded ONUs are always scheduled ahead of heavily loaded ones. If there is more than one lightly loaded ONU at a given cycle n, we assume that their transmission order is based on previous cycle (n – 1) transmission order; that is, an ONU that reports its queue status first is scheduled first. We further assume that the order of heavily loaded ONUs transmission is based on the shortest request first (SRF); that is, ONU with the shortest bandwidth request is scheduled first (in ascending order) and ties are broken by the ONU ID.

Note that, if ONUs transmissions are scheduled in an ascending order, LAN data transmissions assigned to any consecutive time slots are always collision-free. On the contrary, if ONUs transmission are scheduled in a descending order (for instance, ONU5 is scheduled to transmit first and then ONU2), collision between two consecutive time slots is possible (re-circulated LAN data transmitted at the end of the time slot assigned to ONU5 might collide with data transmitted at the beginning of the next scheduled time slot assigned to ONU2). Due to possibility of collision, in this work we will consider transmission in ascending order only.

5.3.3 Performance Evaluation for LAN-DBA

In this section, each ONU generates LAN traffic only. The term “effective load” means the load which is on the link while the “offered load” means the load that is offered to the network. *Offered network load* (ONL) is the sum of the loads offered by each ONU and each ONU contributes to ONL equally by generating the same amount of traffic.

To compare the performance results of the proposed distributed scheme with that of the centralized scheme of [4], we use the same system parameters used therein; a system with 16 ONUs, access link data rate from users to an ONU of 100 Mb/s, and a 1 Gb/s upstream EPON line rate (from an ONU to the OLT). The distance between the OLT and the ONUs varies from 20km to 23km (ring circumference = 3km). Maximum cycle time T_{MAX} = 2ms.

The guard time, T_G, at the centralized architecture, separating two consecutive transmission windows is 5μs [4]. Whereas for the distributed architecture, we set $T_G = 0$, since guard time slots are not needed. Buffer size in each ONU is 10 Mbits.

The traffic model used here is the same as that reported in [4] where each ONU has a number of ON/OFF sources, each with a Pareto distribution governing the lengths of the ON/OFF periods, in order to capture the self-similar nature of Ethernet traffic. Each point on the following plots corresponds to a sample of 100 million packets.

Fig. 5.2 compare the average end-to-end packet delay of LAN traffic as a function of Offered Network Load for both centralized IPACT and distributed DBA schemes. As can be seen from the figure, at all network loads, the average end-to-end packet delays of the distributed scheme is always less than those of the centralized one and at low loads, centralized scheme(IPACT) has about 350-450 μs longer average end-to-end delays than that of the distributed one. This is mainly due to the RTT delay from the ONUs to the OLT (over 210 μs for a 20km trunk). Existence of guard bands at the IPACT also contributes to this delay. In IPACT, we assumed OLT processing time of LAN packets before redirecting them back to ONUs as zero. In the case of the proposed ring-based architecture, the maximum delay between the most distant two ONUs on the ring is about 15 μs.. At higher loads, as queuing delay becomes higher and dominant, the delays get higher and two schemes becomes very similar.

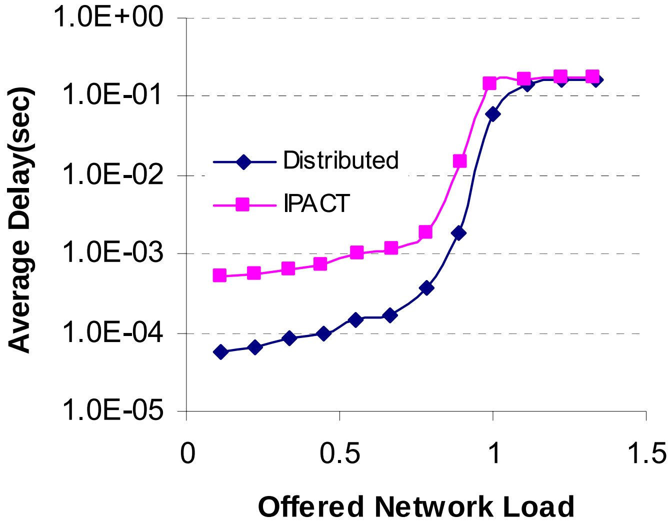


Figure 5.2 Average LAN Packet Delay vs. Offered Network Load

In Fig. 5.3, packet drop rate versus offered network load plot is given. At all loads, especially in light loads, packet drop rate in distributed scheme is less than IPACT. Since at a given cycle heavily load ONUs can get the share of lightly loaded ONUs in distributed algorithm and transmit more packets while in IPACT it can not, this leads to more packets drop in IPACT. Due to T_G presence in IPACT, every subsequent cycles cause more bandwidth wastage. In IPACT, more packets will be dropped from heavily loaded ONUs since heavily loaded ONUs receive heavy traffic but can not send more than B_{MAX}.

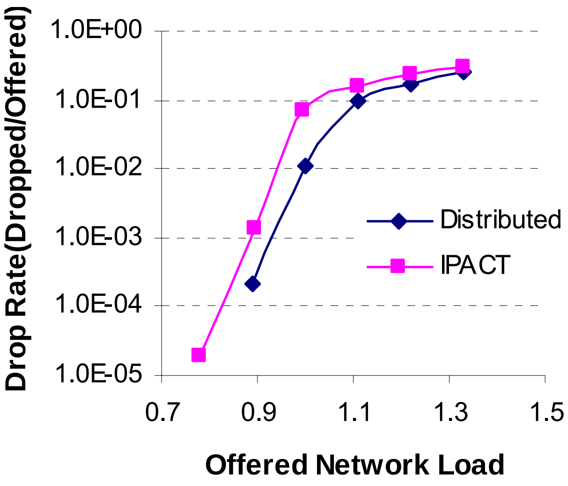


Figure 5.3 LAN Packets Drop Rate vs. Offered Network Load

In Figure 5.4, throughput is same at light loads. At heavy loads, proposed distributed algorithm outperforms IPACT. This happens because in IPACT, between every ONUs transmission there is a guard band which is wasted. So in a given maximum cycle of T_{MAX}, the total amount of bandwidth that is assigned to guard band is equal to T_G * (N-1) where N is the number of ONUs.

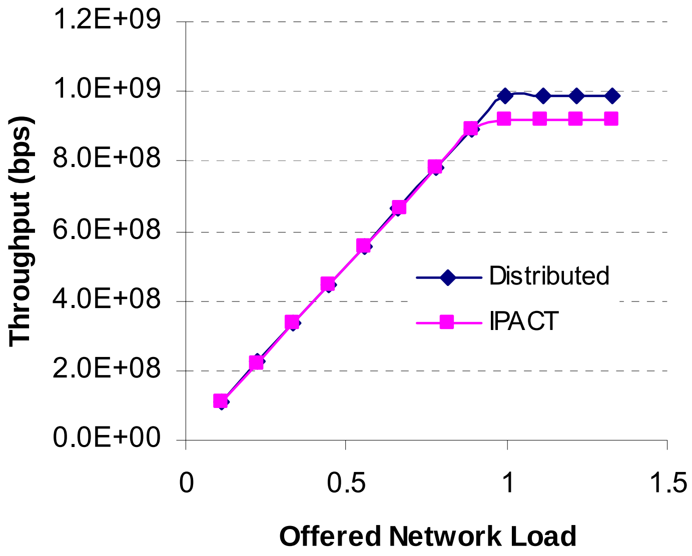


Figure 5.4 LAN Throughput vs. Offered Network Load

5.4 Downstream Wavelength Assignment & Sharing

The OLT houses N downstream queues; each queue, Q_i^{OLT}, is assigned to a specific ONU$_i$ and is connected to a dedicated downstream wavelength, λ_i. Each ONU houses two queues, one queue, $Q_{i,up}^{ONU}$, is assigned to a dedicated upstream wavelength, λ_i, while the other queue, $Q_{i,LAN}^{ONU}$, is assigned to the LAN/control traffic. The process of dynamically assigning/sharing downstream wavelengths is implemented jointly at both the OLT and ONUs as follows: If a dedicated downstream wavelength channel, λ_i, with traffic destined to ONU$_i$ is overloaded (i.e., incoming bursty traffic flows may exceed the dedicated channel rate for some interval, so that its corresponding Q_i^{OLT} is congested), the following steps are executed:

1. The scheduler at the OLT searches for another underutilized/idle downstream wavelength channel, λ_j, i. e., a channel whose corresponding queue, Q_j^{OLT}, has some available space that can accommodate one or more of λ_i's excess flows. This queue will be referred to as an available queue.
2. If the search is successful, the available channel, λ_j, is selected if and only if its corresponding LAN queue at ONU_j, $Q_{j,LAN}^{ONU}$, is also available.
3. The scheduler redirects one, some, or all of λ_i's excess flow(s) from the MAN/WAN to Q_j^{OLT}, where it is then transmitted, along with ONU_j's native downstream traffic arriving from WAN/MAN, to ONU_j over its dedicated wavelength channel λ_j. Such a channel that can accommodate and transport, in addition to its own local downstream traffic, other wavelength channel's downstream traffic, will be referred here as an "Acceptor Wavelength Channel (AWC)" and its corresponding queue as an "available ready queue".
4. The λ_j-downstream optical receiver housed at ONU_j terminates all λ_j's downstream traffic including both native downstream traffic destined to ONU_j and TDS traffic destined to ONU_i, examines the destination MAC address of each detected Ethernet frame, and then performs the following two functions: i) native downstream traffic that matches ONU_j's MAC address is copied and delivered to end-users; ii) TDS traffic destined to ONU_i (whose MAC address does not match that of ONU_j) is redirected to ONU_j's LAN queue and then retransmitted, along with ONU_j's own local LAN traffic, as LAN traffic around the ring over λ_{LAN} to their final destinations, within the proper designated LAN timeslot of ONU_j.

5. TDS flows are returned back to their original queues once these queues are available.

Downstream queues located at the OLT are classified into three different types, a donor queue, an acceptor queue, and a regular queue. A donor queue is defined here as a queue that had some of its own native flows rerouted to other dedicated downstream queue(s). An acceptor queue is a queue that is buffering in addition to its own native local flows, one or more TDS flows that originally belong to other different congested queue(s). A regular queue is neither a donor nor an acceptor queue.

5.4.1 Shared Wavelength Selection & Scheduling (SWS) Algorithm

In this section, a resource allocation scheme that efficiently supports dynamic allocation of wavelengths and sharing traffic among PON end-users is developed. Specifically, we develop a shared wavelength selection/assignment and scheduling (SWS) algorithm that dynamically as well as fairly and efficiently allocates dynamically downstream wavelength channels among end- users. To achieve this objective, downstream network resources (downstream wavelength channels) are load-balanced and efficiently utilized via traffic-engineered routing of end-user's downstream traffic flows. Although the proposed scheme utilizes an OLT-based centralized control plane, its implementation, however, requires the participation of the distributed LAN control plane.

The SWS algorithm is divided into two overlapping processes: The first process, shown in Figure 5.5, is a fixed periodic polling-cycle operation where the flow scheduler at the

OLT periodically checks the status of each dedicated downstream Q_i^{OLT} and its corresponding Q_i^{LAN}. The second, shown in Figure 5.6, is the process of admitting/routing/dropping new native flow(s), which is triggered by the arrival of new native flow(s). The algorithm module maintains two databases. The first database keeps track of the records of all existing and newly arriving flows including flow ID, flow average rate, flow instant rate, native flow's pre-assigned dedicated wavelength channel, and TDS flow's AWC. The second database maintains a Transient Counter (TC_i) for each dedicated downstream queue, Q_i^{OLT}.

The function of a TC is to keep track of the number of native flows that have been redirected from a given donor queue to other available queue(s) as well as the number of non-native TDS flows that are being buffered by an acceptor queue. The number displayed by a TC can either be a zero, positive integer (to indicate that this queue has had a surplus capacity), or a negative integer (to indicate that this queue has had a capacity deficit), corresponding to a regular queue, an acceptor queue, and a donor queue, respectively. For instance, if $TC_i = -2$, this indicate that Q_i^{OLT} is a donor queue that had two of its native flows rerouted to other queue(s). Similarly, if $TC_i = +3$, this indicate that Q_i^{OLT} is an acceptor queue, which is buffering (has buffered) three non-local TDS flows that belong to different queue(s). These databases are updated periodically in each polling cycle upon the arrival of new flows or the termination of an existing flow.

A. First Process: A Periodic Polling-Cycle Operation

This operation is triggered at the end of each polling cycle, where the cycle time is assumed to be fixed ($T_{polling}$). The criterion used to determine whether a downstream queue is congested or available is based on the available size (capacity) of that queue at the instant of polling cycle scheduling, which is equivalent to utilizing a dynamic threshold. In this work, maximum available queue size is set such that worst case packet delay at a given downstream queue does not exceed 10 ms. Thus, for a 1Gb/s output downstream queue rate (all Network DWDM channel capacities are assumed to be equal including downstream, upstream, and LAN channels, each operating at a 1 Gb/s), the maximum available queue size is = (1Gb/s) (10 ms) = 10 Mbits.

If the available size of Q_i^{OLT} at the end of the polling cycle n can't accommodate the accumulated average-sum rate of all existing downstream flows during the next polling cycle n+1, Q_i^{OLT} is congested. In other words, if the following inequality does not hold, Q_i^{OLT} is congested:

$$\{\sum_j f_{j,ave}^{exist} - R_{out}\}\, T_p \leq Q_{i,ava}^{OLT}(t) \qquad (5.8)$$

where $f_{j,ave}^{exist}$ is the average rate of an existing flow j, R_{out} is the downstream output queue rate, T_p is the polling cycle time, and $Q_{i,ava}^{OLT}$ is the available size of queue i at the instant of polling cycle scheduling.

If the inequality of Equation 5.8 holds, this is just an indication that Q_i^{OLT} is not congested but does not necessarily guarantee that it is an available queue. For Q_i^{OLT} to be an available queue, its available size must accommodate the worst case peak-sum rate of all existing flows as well as, at least, one newly arriving flow at its peak rate. Thus, if the following inequality holds, Q_i^{OLT} is an available queue:

$$\{\sum_j f_{j,peak}^{exist} + \sum_K f_{K,peak}^{TDS} - R_{out}\}\, T_p \le Q_{i,ava}^{OLT}(t) \tag{5.9}$$

where $f_{j,peak}^{exist}$ is the peak rate of an existing flow j and $f_{K,peak}^{TDS}$ is the peak rate of a newly arriving flow K or a rerouted TDS flow. To avoid frequent flow transitions that might degrade the performance [6], the average rate of an existing flow in Equation 5.8 has been replaced by the peak rate in Equation 5.9, so that the difference between the available capacity of a non-congested queue and that of an available one is significant.

Similarly, the inequality used to determine the availability of a LAN queue is analogous to that used above in Equation 5.9 for a downstream queue; but appropriately modified to account for all the different input data contents of the LAN queue, including its own local LAN traffic as well as TUS and TDS traffic (if there is any). If the following inequality holds, $Q_{i,LAN}^{ONU}$ is an available queue:

$$\{R_{in,MAX} + \sum_j f_{j,peak}^{exist\,TDS} + \sum_M f_{M,peak}^{exist\,TUS} + \sum_K f_{K,peak}^{new\,TDS} - R_{out}\}\, T_p \le Q_{i,ava}^{ONU}(t) \tag{5.10}$$

where $R_{in,MAX}$ is the line rate from users to ONU_i, $f_{j,peak}^{exist\,TDS}$ is the peak rate of an existing TDS flow j that is utilizing ONU_i, $f_{M,peak}^{exist\,TUS}$ is the peak rate of an existing TUS flow M that is utilizing ONU_i (OLT is aware of both existing TDS flows that it has rerouted itself and TUS flows that it has received via the first component of the LAN signal), $f_{k,peak}^{newTDS}$ is the peak rate of a newly arriving TDS flow K, $R_{out,LAN}$ is the output rate of the LAN queue, and $Q_{i,ava}^{LAN}$ is the estimated available size of the LAN queue calculated at the instant of polling cycle scheduling as follows:

$$Q_{i,ava}^{LAN} = Q_{LAN,MAX} - Q_{Report,\ Aggr} \tag{5.11}$$

It is important to emphasize that in contrast to the actual accurate available size of a downstream queue that is used in Equation 5.9 above, the available size of the LAN queue used in Equation 5.10 is roughly calculated by the SWS algorithm module using Equation 5.11. The SWS algorithm module located at the OLT periodically (each polling cycle) checks the status of each distant ONU's LAN queue ($Q_{Report,\ Aggr}$) via the LAN REPORT messages transmitted to the OLT. Note that the the state of the LAN queues conveyed to the OLT via the REPORT messages might be outdated due to three factors: 1) $T_{polling}$ is not necessarily equal to T_{LAN}; 2) time it takes for the REPORT messages to travel to the OLT (propagation delay between OLT and ONUs; 3) REPORT message transmission time. Thus, the state of the LAN queues used by the SWS algorithm to determine the available size of each distant LAN queue may not reflect the actual instantaneous status of the LAN queues.

At each polling cycle, the algorithm uses the most updated ONU's LAN REPORT message, i. e., the REPORT message that arrives last to the OLT within the current polling cycle. Thus, Equation 5.10 is an approximation and is not as accurate as Equation 5.9 above. However, since Equation 5.10 assumes the worst case scenario for all existing and incoming traffic, it is a sound approximation. This process requires that the fixed polling cycle period, $T_{polling}$, of the SWS algorithm, be at a minimum equal to or greater than the maximum varying length of the DBA algorithm's LAN-cycle, T_{LAN}.

As can be seen from the right part of Figure 5.5, once Q_i^{OLT} has been identified as a congested queue, the status of its corresponding TC_i is checked to determine its type. If Q_i^{OLT} is either a donor or a regular queue ($TC_i \leq 0$), the algorithm determines how many flows should be removed from the queue (rerouted and/or dropped) so that the inequality of Equation 5.8 holds, i. e., Q_i^{OLT} is no longer congested. Note that the actual process of rerouting these excess flows from the congested donor/regular queue to an available ready queue and/or dropping them is postponed until all other remaining queues in the polling cycle are also inspected.

Alternatively, if Q_i^{OLT} is an acceptor queue ($TC_i > 0$), i. e., a queue that is buffering one or more non-local TDS flows, the algorithm first attempts to redirect one or more of these excess TDS flows back to their original source queues (provided that these source queues are available queue) in a LIFO order until the inequality of Equation 5.8 holds. Otherwise, these TDS flows are dropped. Note that in this case, however, an original source queue that restores its own native flows must be an available queue only (i.e., a queue that satisfies only the inequality of Equation 5.9) and need not be an available ready queue, i.e., a queue that satisfies inequalities of both Equations 2 and 3.

The left part of Figure 5.5 shows the steps involved at the end of a given polling cycle n to determine whether a non-congested queue can also be qualified as an "available ready queue"; i.e., a queue, which is ready to buffer, during any instant within the next polling cycle n+1, one or more non-local TDS flows that belong to other queue(s). Once Q_i^{OLT} has been identified as a non-congested queue (satisfies inequality of Equation 5.8), its status is further examined to determine if it qualifies as an available queue (satisfies inequality of Equation 5.9). If it does, the status of its corresponding TC_i is checked to determine its type. If it is either a regular or an acceptor queue ($TC_i \geq 0$), the status of its corresponding LAN queue, $Q_{i,LAN}^{ONU}$, is examined to determine whether it is an available queue as well; if $Q_{i,LAN}^{ONU}$ is an available queue (satisfies inequality of Equation 5.10), Q_i^{OLT} is labeled as an "available ready queue".

Otherwise, if Q_i^{OLT} is an available donor queue ($TC<0$), the algorithm first attempts to restore one or more of its native flows (which have been buffered at other downstream queues), until the queue is no longer available (inequality of Equation 5.9 no longer holds). Thus, if a queue is a strict available donor queue, the process of restoring its own native flows has precedence over that of accepting and buffering non local TDS flows that belong to other queues.

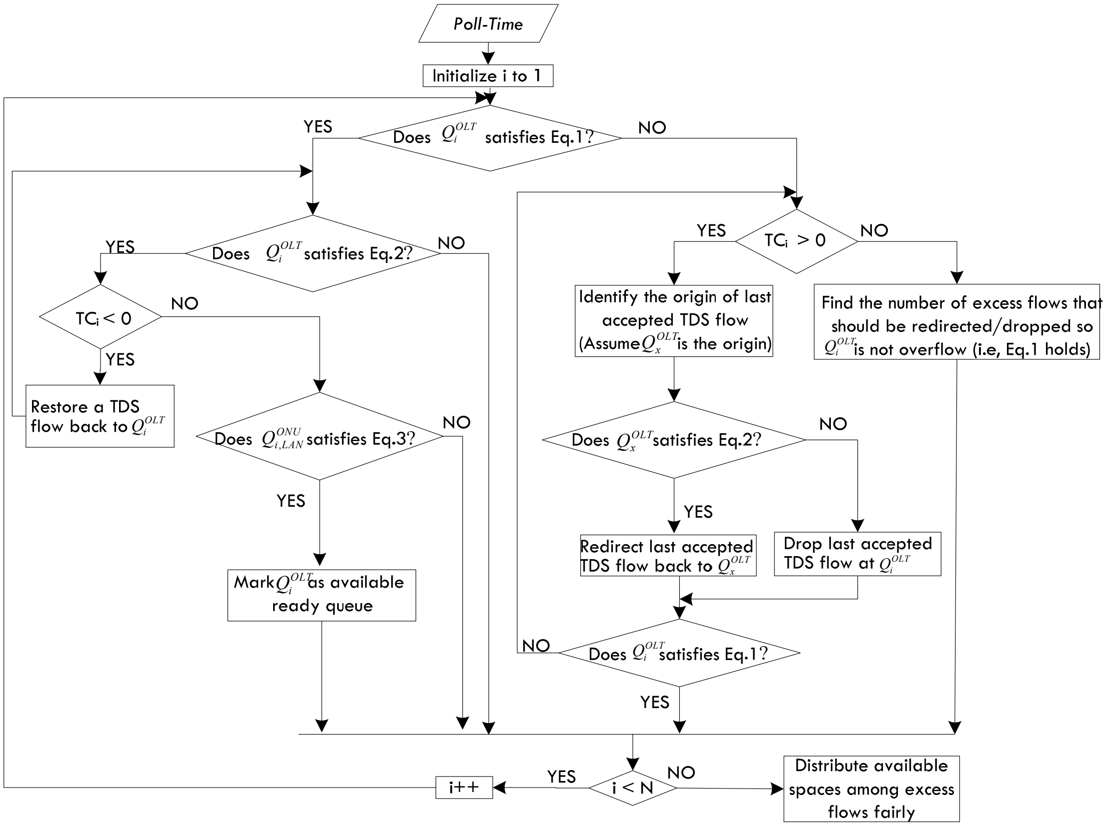


Figure 5.5 Periodic Polling-Time Operations

B. Second Process: Admission & Rerouting of Newly Arriving Native Flows

When a new flow arrives at the OLT from MAN/WAN to the corresponding dedicated downstream queue, as can be seen from figure 5.6, the status of this queue is examined to determine whether it is an available queue or not. If it is (satisfies inequality of Equation 5.9), the flow is admitted and resources are allocated. If Equation 5.9 is not satisfied, i. e., the queue is identified as an "unavailable" queue, flow scheduler further checks whether the queue is an acceptor queue. If it is an acceptor queue, the steps follows is the same as right part of figure 5.5 where the queue is congested and it is an acceptor queue. If the queue is a regular or donor queue, the flow scheduler initiates a quick search to find an available and ready queue. If it finds one, it redirects the newly arriving flow to that queue; otherwise the newly arriving flow would be dropped.

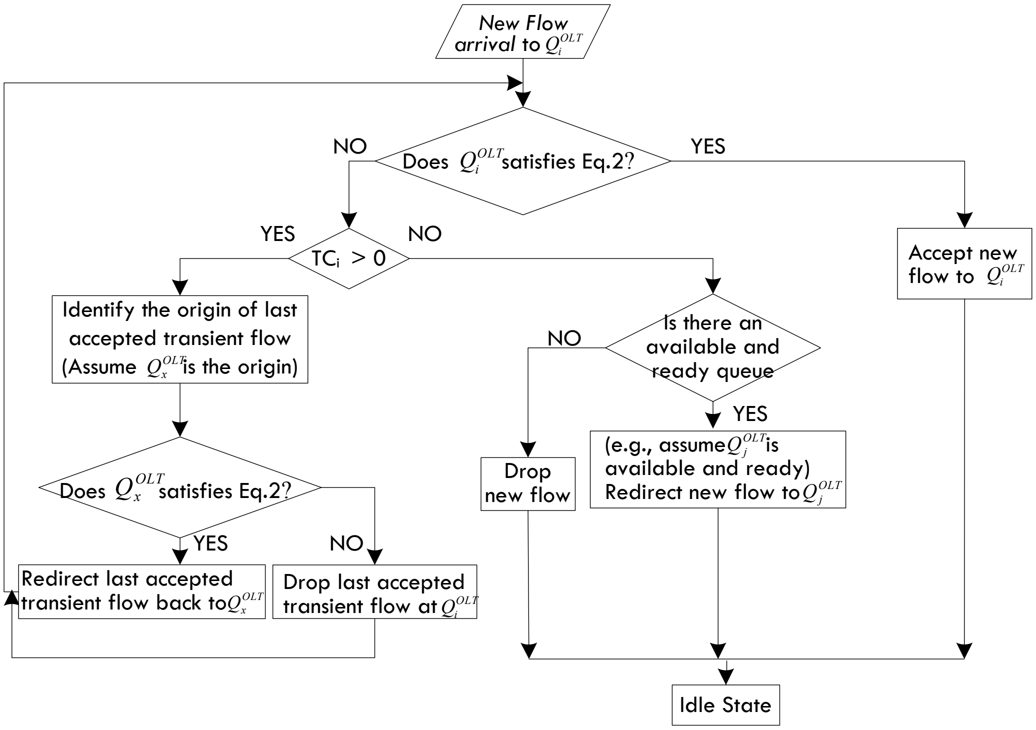


Figure 5.6 Process of admitting/routing/dropping new native flow(s)

C. Traffic-Engineered Selection Criteria

During each polling cycle, as can be seen from Figure 5.5, the algorithm may identify one or more regular/donor congested queues as well as one or more “available ready queues”. Depending on available network resources, a scenario may arise in which not all of the identified congested queues are able to reroute their excess flows to other available ready queues. Another scenario is that there might be only one congested queue but several other available ready queues. Thus, the algorithm, based on some selection criteria, must select which congested queue(s) from amongst all of the identified ones gets to reroute its excess flows; and which available ready queue(s) to be selected from amongst several available ready queues. These criteria must take into account not only maximizing the network throughput, but also being fair to end users; that is, these criteria must balance the trade-off between maximizing throughput and fairness. These selection criteria are:

a) To achieve fairness among end-users, the queue with the least number of exported native flows is selected first. Thus, a regular queue with a zero TC is the queue that is selected to route its excess native flows to other available ready queue(s) first; followed by the donor queue whose TC displays the minimum absolute value, etc. Thus, the donor queue whose TC displays the maximum negative number gets to route its excess native flows last.

b) To maximize downstream throughput, the algorithm selects the least-loaded pair of available downstream and corresponding LAN queues at the OLT and at the ONU, respectively, i. e., the pair with the maximum available average capacity. In both cases, if there is a tie, one is randomly selected.

5.5 Upstream Wavelength Channels Rerouting & Sharing

Analogous to traditional WDM PONs, each ONU in the proposed architecture is assigned a dedicated wavelength for upstream transmission. However, if for a given dedicated user's upstream wavelength channel, $\lambda_{i,up}$, incoming user's upstream bursty traffic flows exceeds the channel rate for some interval, so that its corresponding dedicated upstream queue, $Q_{i,UP}^{ONU}$, is congested. In this case, the flow scheduler at ONU$_i$ may redirect one or more of the user's excess upstream flows to the local LAN queue (these flows are now TUS flows), $Q_{i,LAN}^{ONU}$, provided that this LAN queue has some available space that can accommodate one or more of $\lambda_{i,up}$'s excess flows.

Once these TUS flows are buffered at $Q_{i,LAN}^{ONU}$, they are handled as transient LAN traffic and are transmitted, along with ONU$_i$'s local LAN traffic as well as TDS flows (if there is any) via the LAN/control channel to their final destinations, within the same ONU$_i$'s granted LAN timeslot. Since LAN channel is split into two components at the end of the ring (the first component is destined to the OLT while the second one re-circulates around the ring), the LAN optical receiver housed at the OLT detects the first component of the LAN channel, recovers and processes TUS flows as well as control messages and discards native LAN and TDS traffic (if there is any). On the other hand, either first ONU or source ONU can remove TUS flows from the LAN channel. As it is explained in section 5.3.2, to reduce the idle time on the LAN channel, TUS flows are removed from the link by the first ONU. TUS flows are redirected back to their original upstream queues once they are available.

Our main objective is to simplify the process of redirecting incoming user's excess upstream flows to an available local LAN queue. Since each ONU is directly aware of its own upstream and LAN queue statuses, this process can be directly and independently implemented by each ONU without resorting to any sort of scheduling algorithms. Therefore, the criteria used to determine whether an upstream or a LAN queue is congested or available must be much simpler than those used in the case of a downstream queue. This simple process, which is performed independently by each ONU, is triggered at the end of each LAN cycle (T_{LAN}), is implemented as follows:

- The flow scheduler at each ONU periodically (each LAN-cycle) checks the status of both LAN and upstream queues. Note that the process of periodically checking the status of each LAN queue must be implemented anyway at each LAN cycle since it is required by the LAN DBA module residing at each ONU to perform the LAN timeslots assignment.

- If a given ONU$_i$'s upstream queue, $Q_{i,up}^{ONU}$, is fully loaded, i. e., incoming upstream flows are about to be dropped out, this is an indication that $Q_{i,up}^{ONU}$ is congested, which may trigger the process of redirecting and/or dropping these incoming excess upstream flows. A local LAN queue is an available queue if and only if the following two conditions are satisfied: i) the available capacity of the LAN queue must be higher than or equal to its half maximum size {Q_{LAN} (available) > Q_{LAN} (max)/2; and ii) aggregated incoming traffic rate to Q_{LAN} must be lower than outgoing traffic rate.
- The flow scheduler then checks if the corresponding local LAN queue, $Q_{i,LAN}^{ONU}$, is available. If it is (i. e., satisfies the above two conditions), the flow scheduler redirects one or more of the incoming user's excess upstream flows to $Q_{i,LAN}^{ONU}$, as long as the aggregated incoming upstream traffic rate to $Q_{i,up}^{ONU}$ is lower than outgoing rate. These flows are then transmitted, along with ONU$_i$'s local LAN traffic as well as TDS (if there is any) over λ_{LAN} to their final destinations, within the proper designated LAN timeslot of ONU$_i$. Otherwise, these flows are dropped.
- Once the upstream queue is available, the flow scheduler may return some or all of these TUS flows back to their original source queue, $Q_{i,up}^{ONU}$, from $Q_{i,LAN}^{ONU}$. Analogous to the LAN queue, an upstream queue is an available queue if and only if the following two conditions are satisfied: i) the available capacity of the upstream queue must be higher

than or equal to its half maximum size {Q_{up} (available) > Q_{up} (max)/2; and ii) aggregated incoming traffic rate to Q_{up} must be lower than outgoing traffic rate.

Flowchart pictured as Figure 5.7 shows the Upstream Wavelength Channels routing & sharing algorithm

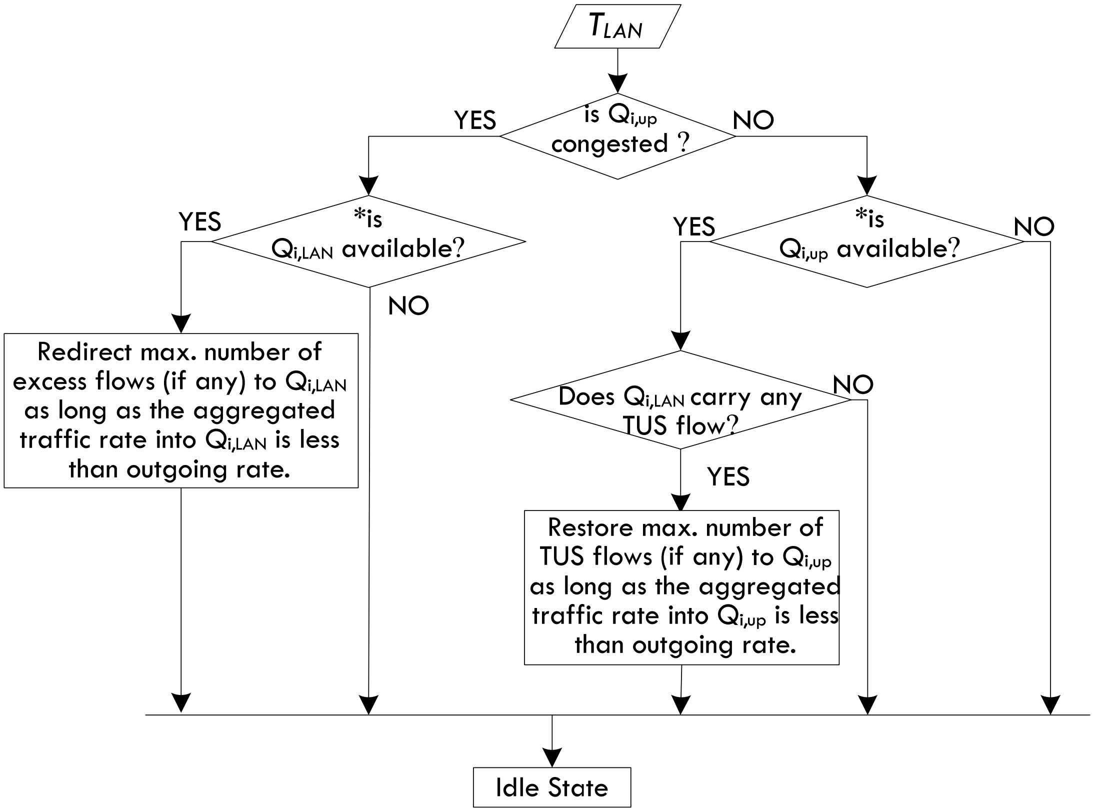


Figure 5.7 Upstream Wavelength Channels routing and sharing

5.6 Power Budget & Scalability of the Proposed Architecture

The scalability of the proposed architecture is mainly limited by the concatenated OADMs insertion losses encountered by both downstream and upstream signals at each node. Since control/LAN signal is regenerated at every node, typical limited power budget problems as well as receiver dynamic range problem (long/short optical network paths and different splitting factors) are totally eliminated.

To examine the performance impact of downstream power budget we consider the worst-case scenario by calculating the total *Optical Distribution Network* (ODN) loss (passive optical elements such as splitters, combiners, fibers, connectors, OADMs and splices forming an optical path) incurred by downstream signal on its longest optical path from the OLT the last ONU. Note that the worst-case scenario ODN incurred by upstream signal on its longest optical path from the first ONU to OLT is almost equivalent to that of the downstream signal.

We assume a 20 Km trunk feeder fiber, first ONU is 20 km away from OLT, and last ONU is 23.2 km away from OLT (ring circumference is about 3.2 km). Table 1 lists best and worst case downstream power budge scenarios for the 1st and 16th ONUs, respectively. All listed parameters, which also include insertion loss, are for commercially available components [13]. Thus, total ODN loss incurred by downstream signal on its path to the first and the last ONU is:

$$L_{Downstream\_Loss}^{Nth\_ONU} = L^{Nth}_{TFF\ MUX} + 2xL_{trunk}^{circulator} + L_{trunk}^{fiber} + L_{Ring}^{fiber} + L_{Mux} + (N-1)xL_{OADM}^{Express} + L_{DROP}^{OADM}$$

(5.13)

For example, if N = 16 ONUs, the total ODN loss using equation 5.12 and 5.13 above (worst case) are 10.5 and 29.1 dB respectively, leading to a 18.6 dB of ODN differential attenuation. Assuming that OLT transmitted power into the fiber is 0 dBm, a downstream receiver with -29.1 dBm sensitivity and 18.6 dB dynamic range is required at the last

ONU to support 16 ONUs. On the other hand, the worst-case scenario ODN incurred by upstream signal on its longest optical path from the first ONU to OLT is almost equivalent to that of the downstream signal = 29.3 dB. The small difference of 0.2 dB arises due to the additional (90:10) coupler loss (0.5 dB), which is encountered by the upstream signal only, while the downstream signal encounters the additional WDM MUX loss (0.3 dB). Upstream and downstream power budget tables are shown in table 5.1 and 5.2

Note that either downstream/upstream power budget demands on receiver parameters are still within a practical and commercial reach. Note also that the 16 nodes limit is not a shortcoming of an architecture that is specifically devised to support a private ring-based local access infrastructure within a 1-4 km diameter area. A typical large private organization would have at most 10-15 buildings within such geographically bounded area. To scale beyond 16 ONUs, a higher transmitted power and/or the use of FEC techniques are required.

COMPONENT		**1st ONU**			**16th ONU**		
		Unit Loss	# of Units	Subtotal	Unit Loss	# of Units	Subtotal
OLT TFF MULTIPLEXER		4	x1	4	1	x1	1
CIRCULATOR		0.6	x2	1.2	0.6	x2	1.2
OADM	EXPRESS CH.	N/A	N/A	N/A	1.4	x15	21
	ADD/DROP CH	1	x1	1	1	x1	1
FIBER LOSS		0.2 dB/km	20 km	4	0.2dB/km	23 km	4.6
RING MULTIPLEXER		0.3	x1	0.3	0.3	x1	0.3
TOTAL(dB)				**10.5**			**29.1**

Table 5.1 Downstream Power Budget for 1st and 16th ONU

COMPONENT		**1st ONU**			**16th ONU**		
		Unit Loss	# of Units	Subtotal	Unit Loss	# of Units	Subtotal
OLT TFF DEMUX.		1	x1	1	4	x1	4
CIRCULATOR		0.6	x2	1.2	0.6	x2	1.2
OADM	EXPRESS CH.	1.4	x15	21	N/A	N/A	N/A
	ADD/DROP CH	1	x1	1	1	x1	1
FIBER LOSS		0.2dB/km	23 km	4.6	0.2 dB/km	20 km	4
SPLITTER(90:10)		0.5	x1	0.5	0.5	x1	0.5
TOTAL**(dB)**				**29.3**			**10.7**

Table 5.2 Upstream Power Budget for 1st and 16th ONU

5.7 Performance Evaluation for Downstream Wavelength Assignment & Sharing

An event-driven WDM-PON simulator was developed using C++. The simulator consists of two modules. The first module executes WDM/TDM-based downstream SWS algorithm at the OLT while the second module concurrently executes TDM-based distributed LAN DBA each ONU. The traffic model used here is the same as that reported in [6-7], where bursty downstream flows are generated by aggregating multiple sub-streams; each is modeled as an on/off source, with Pareto distribution to capture the self-similar nature of Ethernet traffic. Because the execution of the simulation model is computationally intensive and complex, an EPON system with only 8 ONUs is used. The performance metrics used here is average down-link's throughput and delay per ONU.

Based on average down-link traffic load per ONU, ONUs are classified into two different sets, with 4 ONUs each. In the first set, an average down-link load of 0.1(100Mbps) is destined to each ONU from OLT. In the second set, a variable down-link load that is incremented from 0.7 to 1.3 in 0.1 steps. Each ONU in the second will be referred to as "ONU under test" (OUT). Average LAN traffic load is varied from 0.1 to 1.3 in 0.1 steps. Here each ONU contributes to LAN load equally.

The following are the system parameters assumed in the simulation: 1) line rates of all EPON channels including downstream, upstream, and LAN channels, are all equal, each operating at a 1 Gbps; 2) each ONU houses a LAN buffer with 10 Mbits size each; 3) the OLT houses 8 downstream buffers, each corresponds to a given ONU with 10 Mbits size; 4) From MAN/WAN to each downstream buffer, a fixed number of flows (20 flows) arrives constantly where each flow runs at the line rate of 100 Mbps (i.e., Thus, line rate from MAN/WAN to each downstream buffer is 2 Gbps). 6) maximum access LAN link data rate from local users to an ONU is 200 Mb/s; 7) maximum LAN cycle, TLAN = 2 ms; 8) fixed polling cycle, Tp = 2 ms; 9) the distance between OLT and ONUs varies from 20 km to 23 km (ring circumference ≈ 3 km); 10) all downstream traffic is treated as best effort traffic, i. e., all arriving frames at OLT are queued at the corresponding downstream dedicated queues in a first-in-first-out (FIFO) order; 11) strict priority scheduling mechanism is assumed at each LAN queue; that is LAN traffic is assumed to have higher priority compared to both TDS and TUS traffic.

First we have investigated the best value for periodic polling time period. Since LAN cycle is dynamic and maximum LAN cycle period is 2 ms, poll-period can not be less than 2ms (When running downstream SWS algorithm at the OLT, REPORTs should be updated at least once within a poll-period). Based on our model, since 4 lightly loaded downstream channels are available every time to carry TDS traffic, congestion would be more likely at the ONU's LAN buffers (λ_{LAN}). This would be true even in real-life scenarios, since besides native LAN channel, λ_{LAN} might be carrying TDS and TUS traffic. That is why; we have investigated maximum available LAN buffer size at the ONUs when a flow drop is triggered at the OLT. In Figure 5.8, we have investigated maximum available LAN buffer size at an ONU when a flow drop is triggered at the OLT. In Figure 5.8, native LAN load is set to zero. The figure shows that the higher the Tp, the less efficiently LAN buffer is utilized (more available space at the LAN buffer is wasted). This can be easily deduced from equation 5.10, since having higher poll time (Tp) puts more stringent requirement on allowing flows to run as TDS.

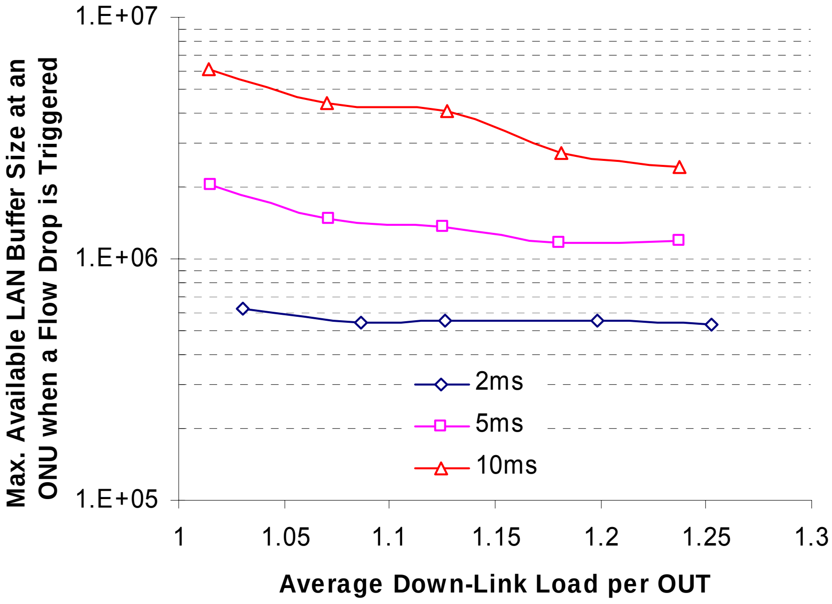


Figure 5.8 Maximum available LAN buffer size at an ONU when a flow drop is triggered at the OLT vs. average Down Link load per OUT

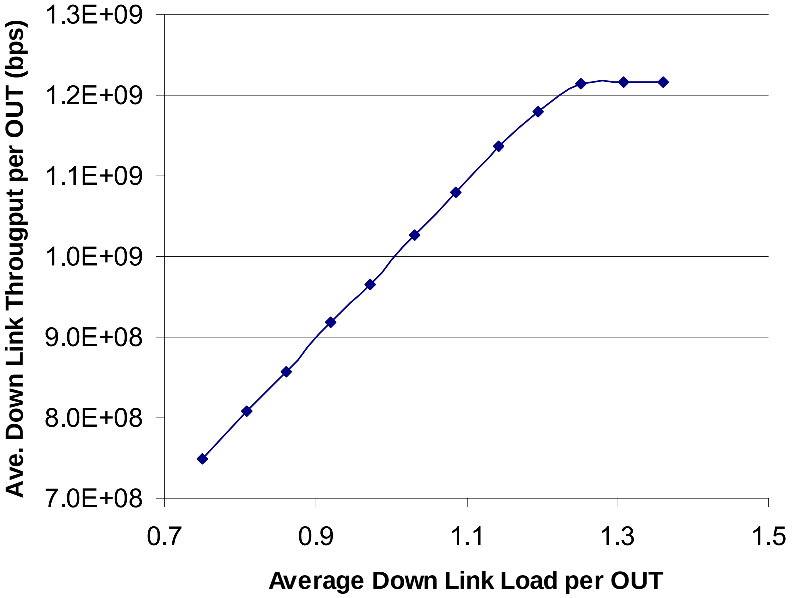


Figure 5.9 Average throughput per OUT versus average down-link load per OUT under zero native LAN traffic

Figure 5.9 shows the down-link throughput per OUT as a function of average down-link load per OUT under zero LAN Load. As can be seen from Figure 5.8, under zero native LAN load and heavy down-link load, an OUT's down-link throughput reaches about 1.22 Gbps. This number is reached due to excess downstream traffic of an OUT is carried as TDS traffic. Note that the theoretical OUT's down-link throughput is given by: {Dedicated downstream channel Rate (1 Gbps) + Unused LAN Capacity}/Number of OUTs (in this case number of OUTs is 4). Thus, under zero native LAN load, the OUT's maximum down-link throughput = 1.25 Gbps, which is in good agreement with the simulation results of Figure 5.9.

Figure 5.10 shows the down-link throughput per OUT as a function of both native LAN and OUT's down-link loads. As LAN load increases, down-link throughput decreases from the peak throughput (1.22 Gbps). Finally, as expected, as the offered LAN load approaches unity or higher, down-link throughput levels-off at around 1 Gbps (dedicated channel rate). This is because transmission of TDS traffic is no longer possible. Note that in traditional WDM-PON systems, an ONU's downstream/upstream rate can not be more than its dedicated rate. Dashed red line in figure 10 shows the throughput per ONU in traditional WDM-PON systems

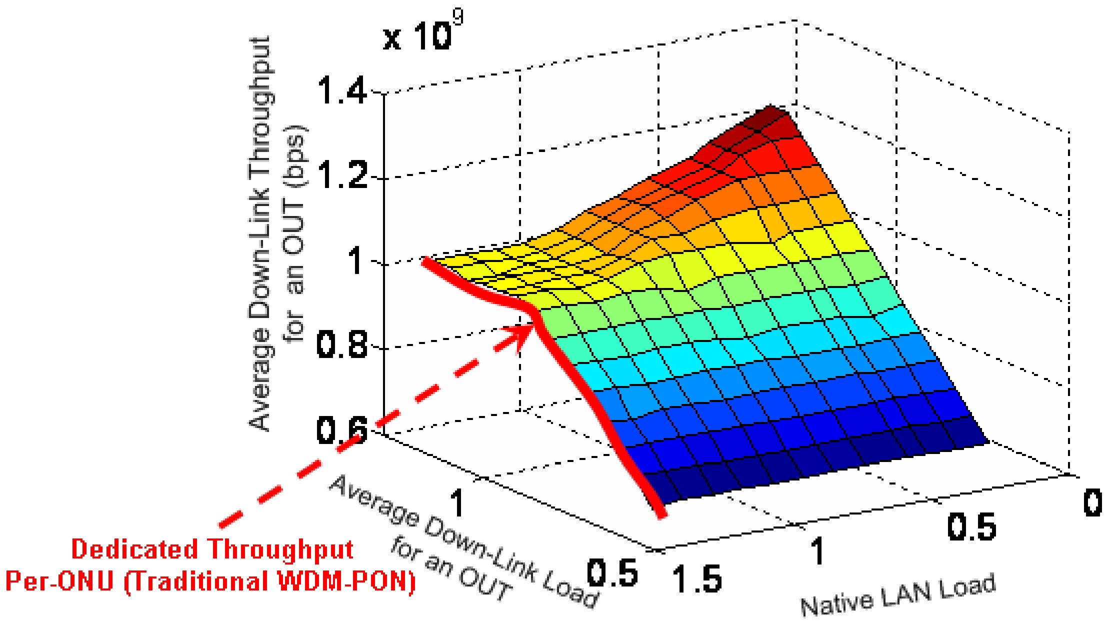


Figure 5.10 Average throughput for an OUT with respect to average down-link load per-OUT and LAN Load

Figure 5.11 shows the average down-link delay versus both native LAN and OUT's down-link loads. At light down-link loads (arrow 1), all downstream flows are transported via their dedicated downstream wavelength channels. In this case, queuing delay is negligible and average down-link delay is mainly due to downstream propagation delay from OLT to ONUs, which is approximately 108 μs. As down-link loads increase, OUTs' downstream buffers start to fill up (shown with arrow 2), average delay gradually starts to increase due to queuing delays. Note that as long as down-link load is light, downstream delay is independent of LAN traffic load.

However, at higher down-link loads where OUTs' downstream queues start to get congested, some downstream flows will be transported as TDS flows via other underutilized wavelength channels (AWCs) whose corresponding dedicated downstream queues are lightly loaded (first set of ONUs). As long as the sum rate of all TDS flows and local LAN traffic is less than 1 Gbps, these TDS flows will not experience additional queuing delays at the LAN queues. These TDS flows will only experience propagation delay, which is composed of two components: 1) propagation delay from OLT to transient ONU (108 μs), and 2) propagation delay across the ring from the transient ONU to the final destination ONU (0.0 μs). Thus, total propagation delay of these TDS flows is almost identical to that of the regular downstream flows (108 μs), since propagation delay of the second component across the ring is almost zero. In this case, since the majority of downstream flows (regular flows) are transported via their dedicated wavelength channels, average queuing delay of all downstream flows (including both regular and TDS flows) is less than 10 ms (maximum queuing delay), which is about 8-9 ms (shown with arrow 3).

On the other hand, when the sum rate of all TDS flows and local LAN traffic is more than 1 Gbps, ONU's LAN queues start to fill up and LAN queuing delay starts to play a role. In this case, average delay of all downstream flows attain the highest value (shown with arrow 4) and is dominated by the effects of both downstream queuing delays at the OLT and LAN queuing delays at the ONUs. At very high native LAN load as well as down-link load where each of down-link and local LAN traffic rate is more than 1 Gbps, no further TDS transmission is allowed. In this case, downstream delay is mainly due to downstream queuing delay alone, which is around 10 ms (shown with arrow 5).

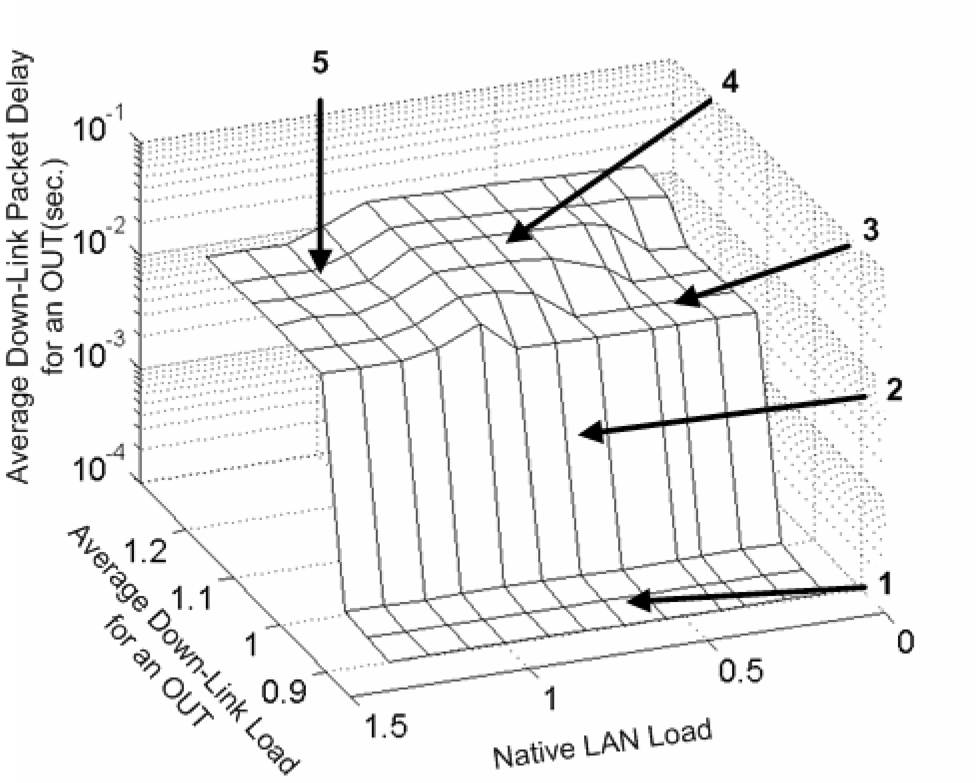


Figure 5.11. Average delay for an OUT with respect to average down-link load of an OUT and LAN Load

5.8 Conclusion

This work has proposed a simple and cost effective local access WDM-PON architecture that combines the salient features of both traditional static WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and high security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end users). Specifically, we have proposed and devised a novel ring-based local access WDM-PON architecture that efficiently supports dynamic allocation of wavelengths /timeslots and sharing traffic as well as a truly shared LAN capability among PON end users.

Chapter 6

Protection & Restoration

6.1 Introduction

Legacy WDM-PON and TDM-PON architectures are deployed as tree topologies, which provide simplicity in the deployment of access networks. However, one of the main limitations of tree-based topologies is the lack of a simple and cost-effective protection and/or restoration capabilities against failures in the access network. The mainstream PON DBA and protection schemes have been centralized – relying on a component at the distant OLT to arbitrate upstream transmission and to detect and recover distribution and trunk fiber breaks [1-9]. The centralized processes of upstream bandwidth allocation as well as detecting and restoring distribution fiber breaks at the distant OLT are lengthy and complex processes and require many changes at each ONU [10].

In the schemes given at [5-9], different automatic protection switching (APS) is used to to handle fiber failures in PON systems. In APS schemes, the optical path of the transported signals is switched to a predetermined path upon detection of either a fiber or equipment failure [11-12]. ITU-T G.983.1 recommended four possible protection schemes, which duplicate fibers and equipment at the ONUs and OLT [8]. These schemes are all centralized and can significantly alter the cost-effectiveness of the PONs since they require many redundant components as well as many spare fibers connections to each ONU [9-10].

This chapter proposes two different self healing architectures for the proposed WDM-PON ring architecture, namely fully distributed self healing WDM-PON ring architecture and hybrid self healing WDM-PON ring architecture. Both architectures provide simple and cost-effective resilience capabilities against any and all kinds of networking failures. Both of the proposed schemes are capable of protecting against both node (ONU) and distribution/trunk fiber failures. These schemes enable the restoration of all network traffic including upstream, downstream, and LAN data. The proposed schemes can also protect against any combination of concurrent double failures including trunk/distribution fiber breaks and node failures.

The simulation results indicate that the performance of the both schemes, including traffic loss and restoration speed outperforms the standard schemes which are recommended by ITU-T G.983.1. Furthermore, the recovery time associated with any and all different distribution network/trunk failures is still within the delay-bound limit required for delivering guaranteed triple play services.

6.2 Fully Distributed Self Healing WDM-PON Ring Architecture

Figure 6.1 illustrates the fully distributed self healing WDM-PON architecture. The solid lines (working fiber) represent the normal state architecture while the dotted lines (protection fiber) represent the redundant protection components. The protected architecture is identical to that of the normal working architecture except for the following additional components (dotted lines): i) a redundant short distribution fiber ring and a trunk fiber; ii) two 2x1 optical switches located at the OLT; iii) Automatic Protection Switching (APS) module located at each ONU;

6.2.1 Protected State Operation

The APS module attached to each ONU is the basic building block of the proposed self-healing mechanism that monitors the state of its adjacent distribution fiber paths and its own state and performs both fault detection and automatic switching process. The APS module connects to both incoming and outgoing working and protection fibers. Each APS module houses a commercially available low loss 4x4 bidirectional Optical Switch (OS) that is capable of switching from any input port to any output port [13]. It also includes two detection circuits, where each circuit comprises a band splitter (to separate the combined downstream/upstream/LAN signal into its constituents LAN and downstream/upstream signals), a control circuit to configure the OS, and a p-i-n detector (except the first detection circuit of ONU_1, which has two p-i-n detectors).

The first/second detection circuit of each ONU (except ONU_1) is used to detect the LAN signal via taping a small portion (about 1%) of the incoming/outgoing combined signal and then passing it through the band splitter. However, the first detection circuit of ONU_1 is used to detect both the LAN and downstream signals via taping a small portion (about 1%) of the incoming combined signal and then passing it through the band splitter. Under normal operation, as can be seen from Figure 1, the combined signal traverses the incoming and outgoing working fibers via ports 2-5 and 8-3, respectively.

In general, we classify failure scenarios into three different classes, a trunk failure, a general distribution link failure and a general node (ONU) failure. A general distribution link is defined here as a fiber segment that connects two adjacent ONUs. All links connecting adjacent ONUs are general distribution links except the following two special links, which require different detection and recovery mechanisms: 1) the distribution fiber segment that connects the first ONU (ONU_1) and the circulator; this link will be referred here as the first link; 2) the distribution fiber segment that connects the last ONU (ONU_N) and the circulator; this link will be referred here as the last link. All nodes distributed around the ring are general nodes except the last ONU, which requires only different recovery mechanisms. As will be shown below, all links and nodes that are at the trunk-ring junction (trunk, first and last links, first and last ONUs) have special significance.

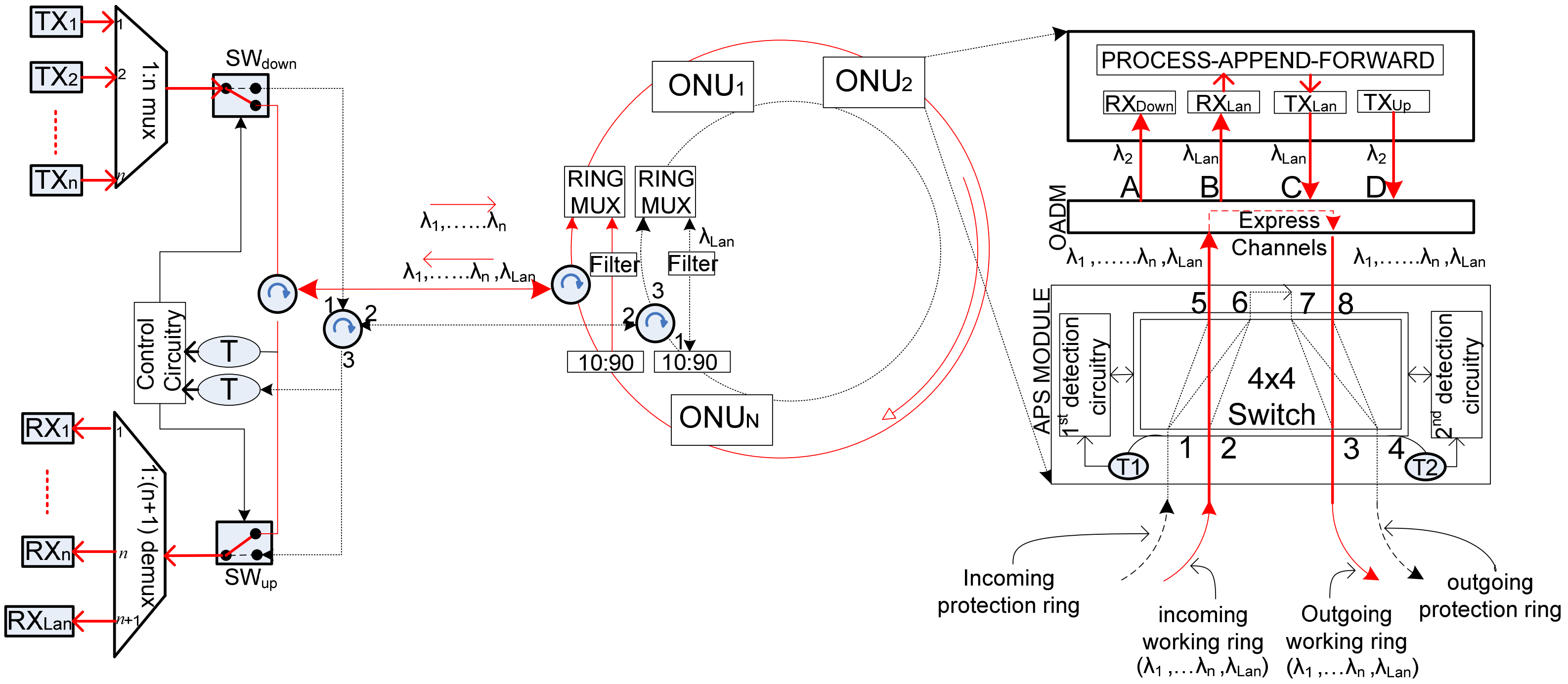


Figure 6.1 Fully Distributed Self Healing WDM-PON Ring Architecture

6.2.2 Synchronization & Resynchronization

Under normal operation, all ONUs are synchronized to a common reference clock extracted from OLT’s downstream traffic. Clocking information, in the form of a synchronization marker, is encapsulated within each ONU’s dedicated downstream wavelength, at the beginning of each downstream frame cycle. The synchronization marker is a one-byte code that is transmitted every 2 ms to synchronize all ONUs with the OLT [2-3]. The TDM controller at each ONU, in conjunction with timing information from the OLT that takes into account the ranging problem (the distance, and hence propagation delay, between the OLT and different ONUs on the ring is not the same), controls the LAN transmission of the variable-length packets within the dedicated time slots. Maintaining proper time sync between different channels is required for the appropriate operation of the LAN DBA algorithm.

In the event of a failure, normal ONUs transmissions stop and the sync between ONUs and OLT is lost. This nullifies the LAN cycle’s timeslots granted to ONUs. Once the failure recovery is complete, a new cycle’s timeslots must be recalculated via the LAN-DBA module located at each ONU. This requires reestablishment of synchronization between ONUs and OLT again. As will be shown below, this process is managed by the affected ONU.

6.3 Fully Distributed Fault Detection and Recovery Mechanisms

In this section, we present fully distributed fault detection and recovery schemes, where the proposed mechanisms are independently initiated and managed by the affected ONU. The ONUs utilize a fully distributed control plane that enables them to directly communicate their status and exchange signaling and control message information with one another. The control plane further enables the ONUs to independently detect, manage, and recover any and all networking failure scenarios except for three special cases, namely, trunk, last link, and last node failures. In each of these three special cases, as will be shown below, both ONU_1 and the OLT jointly participate in the detection and recovery process.

6.3.1 Decentralized Fault Detection Mechanisms

Direct intercommunication among ONUs is implemented via the LAN/control wavelength channel, where the control messages (REPORTs) along with LAN data share the same LAN channel bandwidth (in-band signaling). The proposed scheme utilizes a fully distributed time division multiple access (TDMA) arbitration scheme in which the OLT is excluded from the arbitration process. It assumes a cycle-based LAN link, where the cycle size can be either fixed or variable length confined within certain lower and upper bounds to accommodate the dynamic LAN traffic conditions.

During each LAN cycle, ONUs transmit their REPORT messages along with LAN data sequentially in an ascending order within their granted time slots around the ring from one node to the next, where each REPORT message is finally removed by the source ONU after making one trip around the ring. Since REPORT messages are processed and retransmitted at each node, ONUs can directly communicate their status and exchange signaling and control message information with one another.

Under normal operation, the REPORT message typically contains the desired size of next timeslot based on the current ONU’s LAN buffer occupancy. ONUs exchange REPORT messages concerning their LAN queue status and their transmission needs amongst themselves. Then, the ONUs sequentially and independently run instances of the same LAN-DBA algorithm outputting identical bandwidth allocation results [14]. Once the algorithm is run, the ONUs sequentially and orderly transmit LAN data without any collisions, eliminating the OLT's centralized task of processing requests and generating grants for bandwidth allocations.

In the event of a failure scenario, the REPORT message transmitted by the affected ONU typically contains a failure indication alarm message that includes specific instructions to both the OLT and a remote node that will be involved in the recovery process. Since the LAN signal is always present on the ring and trunk (cyclic control message is always transmitted independent of the presence or absence of LAN data), general failure detection scenarios (general distribution link and node failures) will primarily be based on detecting the absence/presence of the LAN signal only. Thus, all ONUs are continuously monitoring the status of the LAN signal on both incoming and outgoing fibers.

If the first control circuit of a given ONU$_n$ detects the absence of the LAN signal on its incoming working fiber, a general distribution link failure is assumed. This is the link that interconnects ONU$_{n-1}$ with ONU$_n$. On the other hand, each of the three special links (trunk, first and last distribution links) that are the junction of the OLT, ONU$_1$, and ONU$_N$ requires its own different failure detection mechanism. As shown in Table 6.1, all three failure scenarios are detected and managed by the first ONU and each requires monitoring both the LAN and downstream signals. Thus, the first ONU is the only node that is required to monitor both the LAN and downstream signals. The detection and recovery process for each and every link failure is identical, except for the above three special link failures, which require different detection and recovery mechanisms.

Special Fiber Failure	LAN signal	Downstream Signal
Trunk	√	X
Last Link	X	√
First Link	X	X

Table 6.1 ONU1 failure detection table

On the other hand, if the first control circuit of a given ONU$_n$ detects presence of the LAN signal on its incoming working fiber while the second control circuit detects absence of same signal (after being processed, regenerated, and retransmitted by ONU$_n$)

on its outgoing working fiber, a node (ONU_n) failure is assumed. While ONU_n detects its own failure (via an APS module attached to it), however, as will be shown, managing the failure is delegated to the next node on the ring (ONU_{n+1}). Thus, failure of the last ONU is also managed by the first ONU. This means that, overall; the first ONU manages all four special failure scenarios including the three special links plus the last node failures. The detection and recovery process for each and every node on the ring is identical, except for the last node, which requires a different recovery mechanism.

6.3.2 Fully Distributed Recovery Processes

In general, the recovery process is implemented via the participation of three cooperating network nodes including the affected node (ONU_n), OLT, and either ONU_{n-1} (in the case of a link failure) or ONU_{n+1} (in the case of a node failure).

6.3.2.1 General Link Recovery: The successful completion of the recovery process of a given general link failure scenario involves the following steps:

1) To avoid false failure detection, once the affected node (for instance ONU_n) detects a given link failure (e.g., fiber cut), it must wait for a predetermined timeout.

2) ONU_n then performs the following three functions: i) stops both LAN and upstream transmissions; ii) switches to the incoming protection fiber; and iii) floods the network with a failure indication alarm message (first REPORT message) that includes specific instructions to ONU_{n-1} (to switch its transmission from outgoing working fiber to outgoing protection fiber), each and every other ONU (to stop both LAN and upstream transmissions), and OLT (to stop downstream transmission).

3) ONU_n keeps flooding the network with the failure message expecting its failure frame to loop back to it via ONU_{n-1}'s outgoing protection fiber.

4) Upon receiving the failure message, each and every ONU on the ring stops LAN and upstream transmissions; likewise, OLT stops downstream transmission.

5) Once ONU_n receives back its failure frame (assuming ONU_{n-1} has already switched to the outgoing protection fiber), it starts flooding the OLT with a second REPORT message requesting downstream resynchronization frames.

6) Once the OLT receives a resynchronization request from ONU_n, it resumes all downstream transmissions.

7) Once ONU_n receives resynchronization frames from the OLT, it initiates a new cycle (recovery process is now complete) by transmitting its normal REPORT control message to all other ONUs. Then, all ONUs sequentially send their REPORTs; once all reports are exchanged for LAN-DBA calculation of the new cycle, new grants are calculated and normal operation resumes.

6.3.2.2 General Node Recovery: The recovery process of a general node failure involves the following steps:

1) Once the APS module attached to a given node (for instance ONU_n) detects its failure, it waits for a predetermined time period (timeout) and then initiates the process of reconfiguring its OS to the bypass mode by switching the incoming signal directly to the outgoing protection fiber via ports 2-6-7-4 (Figure 6.1). Due to its failure, ONU_n cannot broadcast a failure indication message to its adjacent node (ONU_{n+1}) or to any other node.

2) While ONU_n's APS module is initiating the switching process, ONU_{n+1} detects the absence of LAN signal on its incoming working fiber and erroneously assumes a distribution link failure between itself and ONU_n.

3) ONU_{n+1} then starts the process of a general link recovery (steps listed in Section 6.3.2 above).

4) ONU_{n+1} performs the following three functions: i) stops all transmissions; ii) switches to the incoming protection fiber; and iii) floods the network with a failure indication alarm message that includes specific instructions to ONU_n (to switch its transmission from the outgoing working fiber to the outgoing protection fiber), each and every other ONU (to stop both LAN and upstream transmissions), and OLT (to stop downstream transmission).

5) ONU_{n+1} keeps flooding the network with the failure message waiting to receive back its failure frame via ONU_n's outgoing protection fiber.

6) ONU_n cannot receive or process ONU_{n+1}'s request message. However, ONU_n's control circuit is already reconfiguring the OS to the bypass mode (switching to the outgoing protection fiber). In other words, ONU_n's APS module is indirectly implementing the request message.

7) Once ONU_{n+1} switches to the incoming protection fiber and ONU_n switches to the outgoing protection fiber (i.e., ONU_n is bypassed), ONU_{n+1} receives back its failure message and proceeds with steps 5 and 6 of the general link recovery process listed above.

8) Once ONU_{n+1} receives resynchronization frames from the OLT, it initiates a new cycle by transmitting its normal REPORT message to all other ONUs. Then, all ONUs

sequentially send their REPORTs. In this case, however, due to ONU_n failure, ONU_{n+1} will not receive its REPORT. ONU_{n+1} then recognizes that the absence of LAN signal is in fact due to the failure of ONU_n (and is not due to a link failure) and that ONU_n is now in the bypass state.

9) ONU_{n+1} starts the LAN-DBA calculation for the new cycle without ONU_n's REPORT; new grants are calculated and normal operation resumes.

Two comments are in order here. First, note that when ONU_n fails, it is ONU_{n+1} that undertakes the management of the recovery process assuming an erroneous failure scenario (link failure). Second, note that the OLT's role in the detection and recovery processes of both the general link and node failures has been so far limited to receiving and processing the affected ONU's requests (failure indication and resynchronization messages).

6.3.2.3 Special Failure Scenarios Recovery: The recovery mechanism of each of the four special failure scenarios including the three special links (trunk, first, and last links) as well as last node failure is almost identical and, in each case, requires the participation of all three nodes at the junction where the trunk intersects with the ring fibers (OLT, ONU_1, and ONU_N), where all the involved parties must switch to the protection fiber. Note that it is the first ONU that manages all of these four special failure scenarios. All steps associated with the recovery of a general link/node failure described above are also applicable to these four special link/node failure scenarios except that the OLT's role must be extended now to include switching its transmission to the trunk protection fiber.

There are two options to implement the switching process. The first option is to extend ONU_1's failure indication alarm message to include an additional request to the OLT to switch its transmission to the protection fiber. The second option is that the OLT itself must independently implement the switching process. Note that upon either trunk, last link, or last node failure, the connectivity between ONU_1 and OLT is lost and, therefore, the first option is not applicable. This means that in each of these three special cases the OLT must independently stop downstream transmission and switch to the protection fiber. However, each and every failure scenario of these three cases cuts off the flow of the LAN signal to OLT. Thus, in each case of these three cases, the OLT can independently detect absence of the LAN signal and then automatically stops all downstream transmissions and switches to the protection trunk fiber. Note that the final destination of ONU_1's first REPORT message is now the last ONU.

The fourth special failure scenario, "first link failure", is the only special case where ONU_1 can communicate with the OLT (connectivity is not lost) and, hence, flow of the LAN signal to the OLT is not interrupted. In this case, the second option is no longer applicable and OLT must rely on ONU_1's additional request to switch its transmission to the protection trunk fiber.

6.3.2.3.1 First Link Recovery Process: The recovery process of first link failure, shown by arrow-1 in figure 6.2 involves following steps: 1) The control circuit of ONU_1's APS module detects the absence of both downstream and LAN signals and, thus, identifies the failure as a first link failure (see Table 6.1). 2) ONU_1 then performs the following three functions: i) stops all transmissions; ii) switches to the incoming protection fiber; and iii) floods the network with an extended failure indication alarm message that includes specific instructions to both ONU$_N$ (to switch its transmission from the outgoing working fiber to the outgoing protection fiber) and OLT (to stop downstream transmission and switch to the protection trunk fiber). 3) Once all involved parties (OLT, ONU$_1$ and ONU$_N$) switch to protection fibers, the upstream/downstream signal path with the OLT is restored.

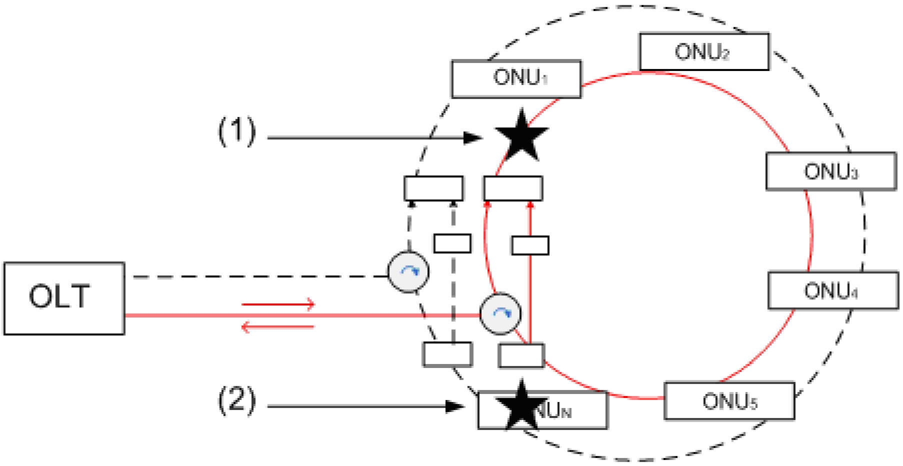


Figure 6.2 First Link Failure (arrow-1), Last Node Failure (arrow-2)

The recovery process of each of the other two special link failures (trunk and last link) is almost identical to that of the first link except for the following: 1) in step 2-iii above, ONU$_1$ floods the network now with rather a short failure indication alarm message that includes no instructions at all to the OLT; 2) the OLT independently detect absence of the LAN signal and then automatically stops all downstream transmissions and switches to the protection trunk fiber.

6.3.2.3.2 Last ONU (ONU_N) Recovery Process: The recovery process of last ONU failure, shown by arrow-2 in figure 6.2 involves following steps The recovery process of this node is identical to that of a last link failure. This is because, as in the case of a general node failure, when ONU_N fails, ONU_1 undertakes the management of the recovery process assuming an erroneous link failure scenario (last link failure). Specifically, when ONU_N fails, ONU_1 detects the absence of the LAN signal and presence of downstream signal and, thus, erroneously assumes that the failure is a last link failure (see Table 6.1). ONU_1 then floods the network with a short failure indication alarm message that includes specific instruction to ONU_N (to switch its transmission from the outgoing working fiber to the outgoing protection fiber). Concurrently, the OLT detects absence of the LAN signal and automatically stops all downstream transmissions and switches to the protection fiber.

Although ONU_N cannot receive and/or process ONU_1's message due to its failure, however, as in the case of a general node failure, ONU_N's control circuit is already reconfiguring the OS to the bypass mode (switching to the outgoing protection fiber). In other words, ONU_N's APS module is indirectly implementing the message.

6.4 Recovery Time Analysis

Recovery time is defined here as the time from when a failure occurs to when service is fully restored and a new cycle resumes. The total recovery time is the sum of the following delay components:

1) Fault detection time, which is assumed here to be ~ 0 sec for all fault detection scenarios that are independently detected via an ONU without OLT's participation. This is true for any and all network failure detection scenarios except for the three special failure cases (trunk, last link, and last node failures) where both ONU_1 and OLT jointly participate in the detection process. In each of these three cases, the detection time is taken as: Max (time it takes for OLT to detect absence of the LAN signal, time it takes for ONU_1 to detect the fault).

Since it takes for ONU_1 ~ 0 sec to detect either the last link or last node failure, while it takes for OLT ~ 100 μs to detect absence of the LAN signal in either case, detection time for either the last link or last node failure is ~ 100 μs. On the other hand, since it takes for ONU_1/OLT approximately the same average time (~ 50 μs) to detect a trunk failure/absence of LAN signal (assuming that the cut occurs at the middle of the trunk), average detection time for a trunk failure is ~ 50 μs.

2) Timeout, $T_{timeout}$, to avoid false fault detection, which is assigned here an arbitrary value of 5.0 μs.

3) REPORT/GATE transmission time, $T_{transmission}$, (64*8 (bits) /1Gbits/s) = 0.5 μs). The recovery process requires two REPORT messages (failure indication message and synchronization request) and one GATE message (synchronization frame).

4) Time it takes for the REPORT/GATE message to travel from the source node/OLT to the OLT/source node. This is the propagation delay, $T_{propagation}$, between the source node and the OLT (worst case scenario when the source node is ONU_1), which has a maximum value of approximately 15 + 100 = 115 μs (corresponding to 3.14 km ring circumference + 20 km trunk feeder). Since there are two REPORT messages from the source node to the OLT and one GATE message from the OLT to the source node, worst case overall REPORT/GATE messages propagation delay ≈ 3 x $T_{propagation}$ = 345 μs.

This is true any and all network failure scenarios except for the three special failure cases listed in item 1 above, where total propagation delay for each of these cases is reduced by about 100 μs. This is because, unlike all general failure scenarios where the first REPORT message sent by the affected ONU has to travel all the way down to the OLT, the first REPORT message, in each of these three special cases, however, only travels to the last ONU. Thus, the first REPORT's propagation delay is now reduced from a 115 μs to ~ 15 μs (corresponding to just one trip around the ring). Thus, in each of these three cases, total propagation delay is now reduced to about 245 μs. Note, however, that the reduction in total propagation delay is almost offset by the corresponding increase in the fault's detection time associated with each of these three cases.

5) REPORT/GATE message processing time, which is assumed here to be ~ 0 sec.

6) One switching time, T_{switch}, (in the case of a single failure), which varies according to the switching fabric.

Thus, the worst case general failure recovery time, $\mathbf{T_{Recovery}}$, is given by:

$$\mathbf{T_{Recovery}} = T_{detection} + T_{timeout} + 3{*}T_{transmission} + 3{*}T_{propagation} + 3{*}T_{process} + T_{switch}$$

$$= 0.0 + 5.0 + 3{*}0.5 + 3\ (15 + 100) + 0.0 + T_{switch}$$
$$= 351.5\ \mu s + T_{switch}$$

The failure recovery time associated with each of the three special failures (trunk, last link, and last node) is almost the same as that of a general failure given above, since the 50-100 μs reduction in total propagation delay is almost offset by the corresponding 100 μs increase in the fault's detection time associated with each. In general, the switching time is much longer than all other delay components combined and, therefore, the total recovery time is mainly dominated by the switching time.

6.4.1 Recovery Process for Concurrent Double Failures

The proposed fault detection and APS recovery mechanisms presented above for a single failure can also be combined to recover from any combination of concurrent double failures including trunk, distribution fiber, and node failures. Similar to the case of a single failure scenario, the recovery time is mainly dominated by T_{switch}. However, depending on the combination of double failures, the recovery time in this case may vary from a minimum of one switching time to a maximum of several switching times. Note, however, that the failure detection times and packet loss associated with concurrent double failures cases are similar to that of single failure cases.

As an illustrative example, Figure 6.3 summarizes a three-phase recovery process for an arbitrary scenario of concurrent double failures {node (ONU_2) and distribution fiber link connecting ONU$_4$ and ONU$_5$, L$_{45}$}, where the recovery time is almost equal to two switching times.

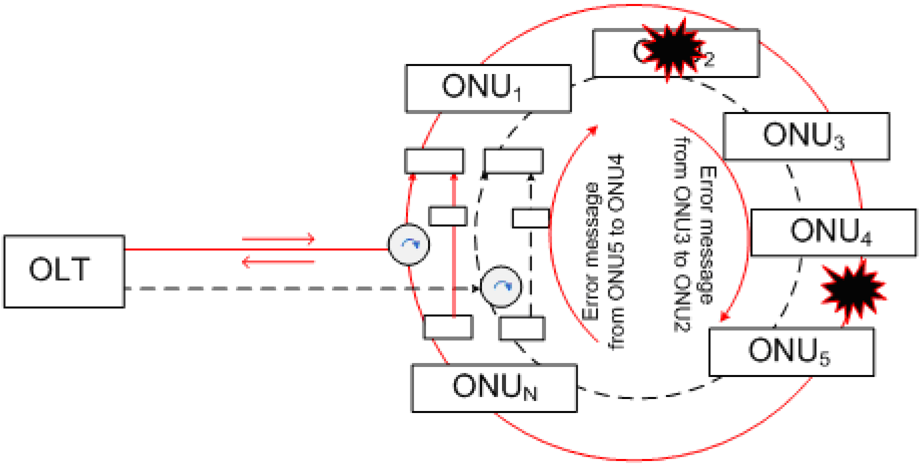


Figure 6.3 Concurrent Failures (Node-2 & Distribution Link-5)

Phase 1

1. APS module attached to ONU$_2$ detects the failure and initiates the process of reconfiguring its OS to the bypass mode by switching the incoming signal directly to the outgoing protection fiber.

2. ONU$_3$ detects absence of the LAN signal on its incoming working fiber and erroneously assumes a distribution link failure between itself and ONU$_2$.

3. ONU$_3$ then starts the process of a general link recovery by switching to its incoming protection fiber.

4. Concurrently, ONU$_5$ detects absence of the LAN signal on its incoming fiber and then: i) switches to the incoming protection fiber; and ii) floods the network with a failure indication alarm message instructing ONU$_4$ to switch its transmission from the outgoing working fiber to the outgoing protection fiber.

5. Both ONU_3 and ONU_5 keep flooding the network with failure messages expecting their failure frame to loop back to them via ONU_2 and ONU_4's outgoing protection fibers, respectively. But these messages will not reach ONU_2/ONU_4 due to L_{45}/ONU_2 failure.

Phase 2 (After T_{switch}):

1. ONU_2 has already switched to its outgoing protection fiber. Almost concurrently, ONU_3 has also switched to its incoming protection fiber. Thus, connectivity between ONU_2 and ONU_3 is now restored leading to the delivery of ONU_5's failure message to ONU_4.
2. ONU_4 starts the process of switching to the outgoing protection fiber.

Phase 3 (After Two T_{switch}):

1. ONU_4 completes the process of switching to its outgoing protection fiber.
2. Connectivity between ONU_4 and ONU_5 is restored. Thus, the optical path between nodes ONU_2, ONU_3, ONU_4, and ONU_5 is reestablished and failure recovery is completed.

6.5 Power Budget and Scalability of the Proposed Distributed Architecture

The scalability of the proposed architecture is mainly limited by the concatenated OADMs insertion loss encountered by both downstream and upstream signals at each node. Compared to the normal state, the signals encounter the additional OS and tap losses at each node. Since the control/LAN signal is regenerated at every node, typical limited power budget problem is totally eliminated.

To examine the performance impact of downstream power budget we consider the worst-case scenario by calculating the total *Optical Distribution Network* (ODN) loss (passive optical elements such as splitters, combiners, fibers, OADMs, OSs, connectors, and splices forming an optical path) incurred by downstream signal on its longest optical path from the OLT the last ONU. Note that the worst-case scenario ODN incurred by upstream signal on its longest optical path from the first ONU to OLT is almost equivalent to that of the downstream signal.

COMPONENT LOSS		**1st ONU**			**16th ONU**		
		Unit Loss	# of Units	Subtotal	Unit Loss	# of Units	Subtotal
OLT TFF MUX.		4	x1	4	1	x1	1
CIRCULATOR		0.6	x2	1.2	0.6	x2	1.2
OADM	EXPRESS CH.	N/A	N/A	N/A	1.4	x15	21
	ADD/**DROP** CH	1	x1	1	1	x1	1
FIBER LOSS		0.2 dB/km	20 km	4	0.2dB/km	23 km	4.6
RING MULTIPLEXER		0.3	x1	0.3	0.3	x1	0.3
OPTICAL SWITCH		0.3	x2	0.6	0.3	x32	9.6
TOTAL(dB)				**11.1**			**38.7**

Table 6.2 Downstream Power Budget Table for 1st ONU and 16th ONU for Fully Distributed Self Healing WDM PON Ring Architecture

We assume a 20 Km trunk feeder fiber, first ONU is 20 km away from OLT, and last ONU is 23.2 km away from OLT (ring circumference is about 3.2 km). Table 6.2 lists best and worst case downstream power budge scenarios for the 1st and 16th ONUs, respectively. All listed parameters, which also include insertion loss, are for commercially available components [13]. Total ODN loss incurred by downstream signal on its path to the last ONU is given by:

$$L_{Downstream\_Loss}^{Nth\_ONU} = L_{OLT\ MUX}^{Nth} + 2xL_{trunk}^{circulator} + L_{trunk}^{fiber} + L_{Ring}^{fiber} + L_{Ring\ Mux} + (N-1)xL_{OADM}^{Express} + L_{DROP}^{OADM} + 2N*OS \tag{6.1}$$

For instance, if N = 10 ONUs, the total ODN loss using equation 1 above (worst case) is ~ 26.7 dB. Assuming that the OLT transmitted power into the fiber is 3 dBm per channel and allowing a 2dB power margin, a downstream receiver with - 25.8 dBm sensitivity is required at the last ONU to support 10 ONUs. On the other hand, the worst-case scenario ODN incurred by upstream signal on its longest optical path from the first ONU to OLT is almost equivalent to that of the downstream signal = 26.9 dB. Note that either downstream/upstream power budget demand on receiver parameters is still within a practical and commercial reach to support 10 ONUs.

On the other hand, if N = 16 ONUs, the total ODN loss using Equation 2 above (worst case) is ~ 38.7 dB. To accommodate such a high loss budget, commercially available erbium doped fiber amplifier (EDFA) can be used at the OLT to amplify the WDM downstream signal. Thus, assuming an EDFA with 15 dB average gain per channel, an OLT transmitted power into the fiber of 0 dBm per channel, and allowing a 2dB power margin, a downstream receiver with - 25.7 dBm sensitivity is required at the last ONU to support 16 ONUs.

Note that downstream/upstream power budget is still within a practical and commercial reach to support 10-16 ONUs. Note also that the 10-16 node limits is not a shortcoming of an architecture that is specifically designed to support a private ring-based local access infrastructure within a 1-5 km diameter area. A typical large private organization would have at most 10-15 buildings within such geographically bounded area. To scale beyond 16 ONUs, higher transmitted power, avalanche photodetectors, and/or the use of FEC techniques are required.

6.6 Hybrid Self Healing WDM-PON Ring Architecture

Figure 6.4 illustrates the proposed self healing ring-based WDM PON architecture. The solid lines represent the normal state architecture while the dotted lines represent the redundant protection components. The protected architecture is identical to that of the normal working architecture except for the following additional components: i) a redundant short distribution fiber ring and a trunk fiber; ii) two 1x2 OSs located at the OLT; iii) an APS module located at each ONU; iv) a redundant OADM located at each ONU.

6.6.1 Protected State Operation

The APS module attached to both the working and protection OADMs is the basic building block of the proposed self-healing mechanisms that monitors the state of its adjacent distribution fiber paths as well as the node to which it is connected and performs both fault detection and automatic switching processes. As can be seen from Figure 1, each APS module houses four Optical Switches (OSs); two of which are 2x2 (switches B and C) and the other two are 2x1 (switches A and D). Each OS can be connected directly to one of the four add/drop ports of either the working OADM or the protection OADM. Note that port # 4 of switch B is permanently connected to port # 4 of switch C.

The APS module also houses two detection circuits, where each circuit comprises a control circuit to configure the OSs, and a p-i-n detector (except the first detection circuit of ONU_1, which has two p-i-n detectors). The first detection circuit is used to detect the dropped LAN signal via taping a small portion (about 1%) of the incoming LAN signal. Likewise, the second detection circuit is used to detect the added LAN via taping a small portion (about 1%) of the outgoing LAN signal. The first ONU's detection circuit is used to detect both the dropped LAN and corresponding downstream signals via taping a small portion (about 1%) of each incoming signal.

As can be seen from Figure 6.4, under normal operation, each ONU attaches to the ring via the input port of the working OADM (this port attaches to the incoming working fiber, point w_{in} in Figure 6.4) and can transmit data onto the ring via the output port of the working OADM, which in turn attaches to the outgoing working fiber (point w_{out} in Figure 1). Each OS is connected directly to one of the four add/drop ports of the working OADM, i.e., in the normal state, ports 1 and 3 of each switch is always connected.

Under the protected state, each ONU attaches to the ring via the input port of the protection OADM (this port attaches to the incoming protection fiber, point p_{in} in Figure 6.4) and can transmit data onto the ring via the output port of the protection OADM, which in turn attaches to the outgoing protection fiber (point p_{out} in Figure 6.4). Each OS is connected directly to one of the four add/drop ports of the protection OADM, i.e., in the protected state, except in the case of a node failure, ports 2 and 3 of each switch is always connected.

In general, we classify failure scenarios into three different classes, a trunk link failure, a general distribution link failure, and a general node (ONU) failure. A general distribution link is defined here as a fiber segment that connects two adjacent ONUs. All links connecting adjacent ONUs are general distribution links except the following two special links, which require only different detection mechanisms: 1) the distribution fiber segment that connects the first ONU (ONU_1) and the circulator; this link will be referred here as the first link; 2) the distribution fiber segment that connects the last ONU (ONU_N) and the circulator; this link will be referred here as the last link. All nodes distributed around the ring are general nodes except the last ONU, which has special significance. As will be shown below, all links and nodes that are at the trunk-ring junction (trunk, first and last links, first and last ONUs) have special significance.

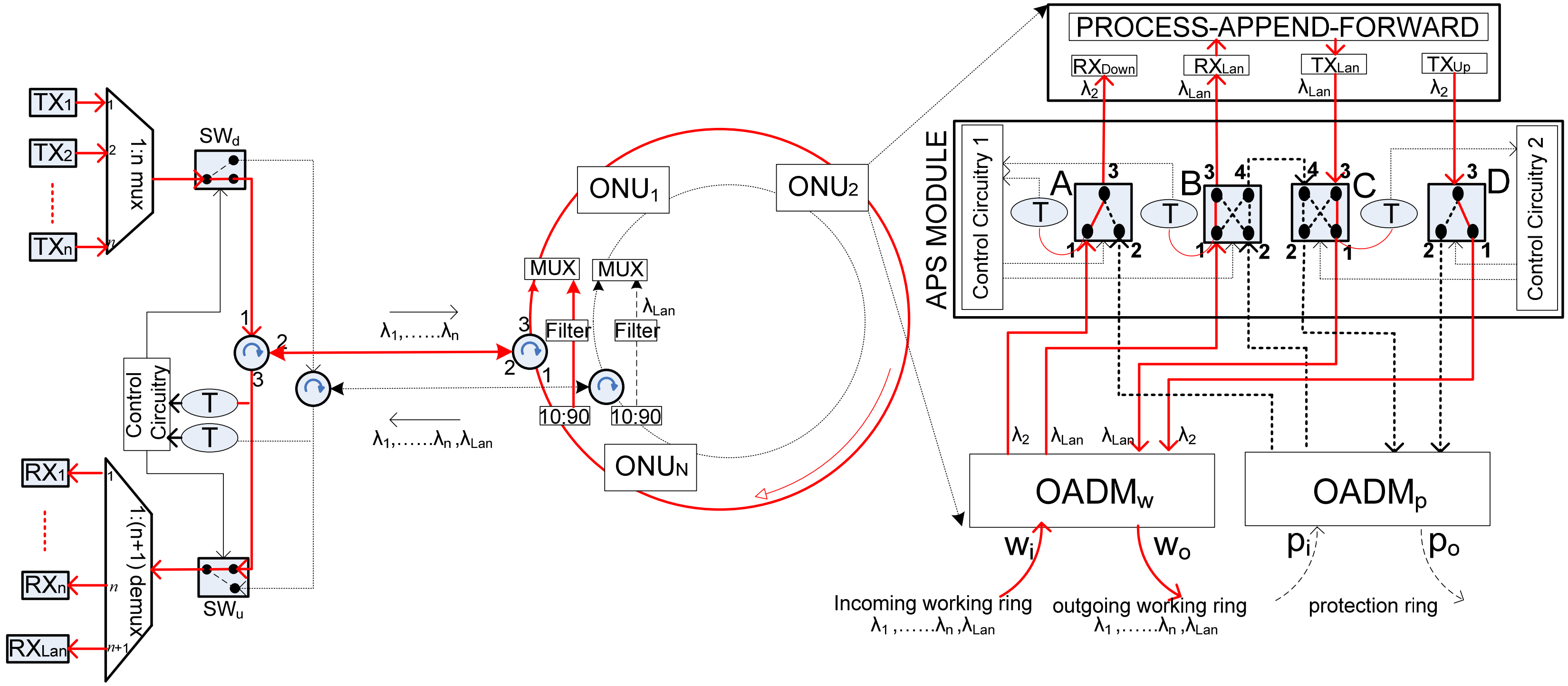


Figure 6.4 Hybrid Self Healing WDM-PON Ring Architecture

6.6.2 Hybrid Fault Detection and Recovery Mechanisms

In this section, we present hybrid fault detection and recovery schemes where in this case both the ONUs and OLT jointly participate in the detection and subsequent recovery process. As will be shown below, the main distinction between the distributed recovery schemes presented above and those of the hybrid schemes are:

1. Under the hybrid approach, the distributed control plane running among the ONUs only enable them to independently detect and manage (but not to fully recover as it is the case in the distributed schemes) all networking failure scenarios, except for the three special cases (trunk, last link, and last node failures). Analogous to the distributed approach, the OLT that must independently detect the failure in each of these three special cases.
2. Unlike the distributed approach presented above where the OLT switches to the protection fiber only during the three special failure cases, the successful completion of each and every hybrid recovery process requires the OLT to switch its transmission to the protection fiber.

6.6.2.1 Hybrid Fault Detection Mechanisms

Analogous to the distributed approach, the ONUs utilize same distributed control plane to directly communicate their status and exchange signaling and control message information with one another. Normal operation of the hybrid approach is also identical to that of the distributed approach presented above.

In the event of a failure scenario, the REPORT message flooded by the affected ONU typically contains a failure indication alarm message with specific instructions to the

OLT as well as all ONUs. Analogous to the distributed approach, since the LAN signal is always present on the ring and the trunk (cyclic control message is always transmitted independent of the presence or absence of LAN data), general failure detection scenarios (general distribution link and node failures) will primarily be based on detecting absence/presence of the LAN signal only. Thus, all ONUs are continuously monitoring the status of the LAN signal on both incoming and outgoing fibers.

If the first control circuit of a given ONU_n detects absence of the dropped LAN signal on the working OADM's LAN drop port, a general distribution link failure is assumed. This is the link that interconnects ONU_{n-1} and ONU_n. On the other hand, each of the three special links (trunk link, first and last distribution links) that are the junction of OLT, ONU_1, and ONU_N requires its own different failure detection mechanism. As shown in table 6.1, all three failure scenario are detected and managed by the first ONU and each requires monitoring both the dropped LAN and corresponding downstream signals. Thus, the first ONU is the only node that is required to monitor both dropped LAN and downstream signals. Though general link's failure detection mechanism is different than that of each of the three special links, the recovery processes, however, are almost identical for both general and special link failures.

On the other hand, if the first control circuit of a given ONU_n detects presence of the dropped LAN signal on the working OADM's LAN drop port, while the second control circuit detects absence of same added signal (after being terminated, regenerated, and retransmitted by ONU_n) on the working OADM's LAN add port, a node (ONU_n) failure is assumed. Analogous to the distributed approach, managing ONU_n failure is delegated to the next node on the ring (ONU_{n+1}). Thus, the first ONU manages all four special failure scenarios including the three special links plus last node failures. The detection and recovery process for each and every node on the ring is identical. Though, as will be shown below, the OLT plays a more active role in the case of three special failure cases (trunk, last link and last node failures).

6.6.3 Hybrid Recovery Processes

In general, the recovery process is implemented via the participation of all network nodes including the OLT and all ONUs.

6.6.3.1 General Link Recovery

The successful completion of the recovery process of a given general link failure scenario involves the following steps:

1) To avoid false failure detection, once the affected node (for instance ONU_n) detects a given fiber break, it must wait for a predetermined timeout.

2) ONU_n then performs the following three functions: i) stops both LAN and upstream transmissions; ii) floods the network with a failure indication alarm message (first REPORT message) that includes specific instructions to each ONU (to stop both LAN and upstream transmissions and to switch to the protection fiber) and the OLT (to stop downstream transmission and switch its transmission to the protection trunk fiber); and iii) switches to the protection fiber (via reconfiguring the four OSs such that each OS is now directly connected to the protection OADM, i.e., ports 2 and 3 of each switch is now connected).

3) ONU_n keeps flooding the network with the failure message expecting its failure frame to loop back to it via ONU_{n-1}'s outgoing protection fiber.

4) Upon receiving the failure message, each and every ONU stops all transmissions and then initiates the process of switching to the protection fiber via reconfiguring the four OSs housed in its APS module; likewise, the OLT, upon receiving the failure message, stops downstream transmission and then initiates the process of switching to the protection trunk fiber.

5) Once ONU_n receives back its failure frame (assuming all ONUs including ONU_n as well as OLT have already switched to protection fiber), it starts flooding the OLT with a second REPORT message requesting downstream resynchronization frames.

6) Once OLT receives resynchronization request from ONU_n, it sends RESYNC frames to each ONU via its dedicated downstream channel, and then resumes downstream transmissions.

7) Once ONU_n receives resynchronization frames from the OLT, it initiates a new cycle (recovery process is now complete) by transmitting its normal REPORT control message to all other ONUs. Then, all ONUs sequentially send their REPORTs; once all REPORT messages are exchanged for LAN-DBA calculation of the new cycle, new grants are calculated and normal operation resumes.

6.6.3.2 General Node (ONU) Recovery

The recovery process of a general node failure involves the following steps:

1) Once the APS module attached to a given node (for instance ONU_n) detects its failure, it waits for a predetermined time period (timeout) and then initiates the process of

reconfiguring its OSs to the bypass mode to switch the incoming/dropped LAN signal (via the protection OADM) directly to the outgoing protection fiber (thus bypassing ONU$_n$) via ports 2-4-4-2 of OSs B and C, respectively. Due to its failure, ONU$_n$ can't broadcast a failure indication message to its adjacent node (ONU$_{n+1}$) or to any other node.

2) While ONU$_n$'s APS module is initiating the switching process, ONU$_{n+1}$ detects absence of the LAN signal on its incoming working fiber and erroneously assumes a distribution fiber break between itself and ONU$_n$.

3) ONU$_{n+1}$ then starts to implement steps 1-2 of the general link recovery listed above.

4) ONU$_{n+1}$ keeps flooding the network with the failure message expecting its failure frame to loop back to it via ONU$_n$'s outgoing protection fiber.

5) ONU$_n$ can't receive or process ONU$_{n+1}$'s request message. However, ONU$_n$'s APS module is already reconfiguring its OSs to the bypass mode.

6) Once ONU$_{n+1}$ receives back its control frame (assuming that ONU$_n$ has already switched to the bypass mode and that all remaining ONUs as well as OLT have already switched to protection fiber), it starts flooding the OLT with a second REPORT message requesting downstream resynchronization frames.

7) Once ONU$_{n+1}$ receive resynchronization frames from OLT, it initiates a new cycle by transmitting its normal REPORT to all other ONUs. Then, all ONUs sequentially send their REPORTs. In this case, however, due to ONU$_n$ failure, ONU$_{n+1}$ will not receive its REPORT. ONU$_{n+1}$ then realizes that the absence of the LAN signal is in fact due to ONU$_n$ failure (and is not due to a fiber break) and that ONU$_n$ is now in bypass state.

8) ONU_{n+1} starts LAN DBA calculation for the new cycle without ONU_n's REPORT; new grants are calculated and normal operation resumes.

Note that upon a node failure, once the OLT has been informed, TDS flows that are utilizing the failed node's LAN queue (if there are any) are either redirected back to their dedicated downstream queues (if possible) or dropped by the flow scheduler at the OLT.

6.6.3.3 Special Failure Scenarios Recovery

As indicated above, the recovery mechanism of each of the three special link failures including trunk, first, and last links is almost identical to that of a general link failure described above. In addition, the recovery process of the last node failure is almost identical to that of a general node failure presented above. Thus, under the hybrid approach, there are only two different standard recovery schemes, one for all link failures and the other one for all node failures including both general and special links and nodes. Though the steps implemented in each of these standard recovery schemes are identical for either general or especial link/node failure, there is, however, a slight distinction between general and special failure (trunk, last link, and last node) recovery only in terms of the mechanism that drives the OLT switching process.

Analogous to the distributed approach, upon trunk, last link, or last node failure, the connectivity between ONU_1 and OLT is lost. Therefore, rather than relying on control instructions (first REPORT message) typically received from the affected ONUs in all other general failure scenarios, the OLT, in each of these three special failure scenario, must independently detect absence of the LAN signal and switches to the protection fiber. As in the case of the distributed approach, each and every failure scenario of these three cases cuts off the flow of the LAN signal to OLT. Thus, in each case of these three cases, the OLT can independently detect absence of the LAN signal and then automatically stops all downstream transmissions and switches to the protection trunk fiber.

6.6.4 Recovery Time Analysis

The recovery time analysis of the hybrid approach is identical to that of the distributed approach and, in general, the total recovery time is mainly dominated by the switching time. However, in the case of concurrent double failures, in contrast to the distributed approach where the recovery time, depending on the combination of double failures, vary from a minimum of one switching time to a maximum of several switching times; the recovery time of the hybrid approach is independent of the combination of double failures, and always equal to just one switching time.

6.6.5 Power Budget & Scalability of the Proposed Architecture

The scalability of the proposed architecture is mainly limited by the concatenated OADMs insertion loss encountered by both downstream and upstream signals at each node. Compared to the normal state, the signals encounter almost no additional loss except for a negligible 0.6 dB loss associated with the OS located at the OLT and an ONU. Since the control/LAN signal is regenerated at every node, typical limited power budget problem is totally eliminated.

Analogous to the distributed approach, the performance impact of downstream power budget is examined by considering the worst-case scenario via calculating the total *Optical Distribution Network* (ODN) loss incurred by downstream signal on its longest optical path from the OLT the last ONU. Note that the worst-case scenario ODN loss incurred by upstream signal on its longest optical path from the first ONU to OLT is almost equivalent to that of the downstream signal.

We assume a 20 Km trunk feeder fiber, first ONU is 20 km away from OLT, and last ONU is 23.2 km away from OLT (ring circumference is about 3.2 km). Table 6.2 lists best and worst case downstream power budge scenarios for the 1st and 16th ONUs, respectively. All listed parameters, which also include insertion loss, are for commercially available components [13]. Total ODN loss incurred by downstream signal on its path to the last ONU is given by:

$$L^{Nth\_ONU}_{Downstream\_Loss} = L^{Nth}_{OLT\,MUX} + 2xL^{circulator}_{trunk} + L^{fiber}_{trunk} + L^{fiber}_{Ring} + L_{Ring\,Mux} + (N-1)xL^{Express}_{OADM} + L^{OADM}_{DROP} + 2*OS \quad (6.2)$$

For example, if N = 16 ONUs, the total ODN loss using the above equation (worst case) is 29.7 dB. Assuming that the OLT transmitted power into the fiber is 3 dBm per channel and allowing a 2dB power margin, a downstream receiver with -28.7 dBm sensitivity is required at the last ONU to support 16 ONUs. Note that either downstream/upstream power budget demand on receiver parameters is still within a practical and commercial reach to support 16 ONUs. Note also that the power budget of the hybrid protection architecture is almost identical to that of the normal state.

6.6.6 Simulation Results

In this section, we assess the feasibility of the proposed distributed detection and recovery schemes. An event-driven packet-based simulation model was developed using C++. The traffic model used here is the same as that reported in [2, 17] where each ONU is modeled as an on/off source, with Pareto distribution to capture the self-similar nature of Ethernet traffic. The performance metric used here is the average traffic loss for upstream, downstream, and LAN traffic. To simplify the analysis, it is assumed that downstream and upstream traffic may utilize only dedicated wavelength channels and that dynamic wavelength sharing is not allowed. This assumption only reduces the computational complexity but doesn't alter the qualitative analysis and/or the general conclusions.

The following are the system parameters assumed in the simulation: 1) an WDM-PON system with 16 ONUs; 2) line rates of all channels including downstream, upstream, and LAN channels are all equal, each operating at 1 Gbps; 3) each ONU houses two queues, a LAN queue and a dedicated upstream queue, each with 10 Mbits size; 4) OLT houses 16 queues, each corresponds to a given ONU with 10 Mbits size; 5) LAN link data rate from users to an ONU is 100 Mbps; 6) downstream link data rate from backbone to each dedicated queue at the OLT is 2Gbps; 7) access upstream link data rate from users to an ONU is 2 Gbps; 8) maximum LAN cycle, $T_{LAN} = 2$ ms; 9) the distance between OLT and ONUs varies from 20 km to 23 km (ring circumference ≈ 3 km); 9) IEEE 802.3ah MPCP REPORT/GATE message is 64 bytes; 10) the LAN DBA scheme used here is the same as that reported in [17].

All network downstream, upstream, and LAN traffic are treated as best effort traffic, i. e., all arriving frames at OLT and ONUs are queued at the corresponding downstream/upstream /LAN queues in a first-in-first-out (FIFO) order. For simplicity, we assume that all ONUs have equal average traffic load including downstream, upstream, and LAN traffic loads. We randomly generate various types of failures and measure the corresponding upstream, downstream and LAN traffic loss. Each point in the simulation results corresponds to a sample of 100 million packets averaged over three different runs.

Note that the number of downstream/upstream WDM channels that experience traffic loss depends on both the type and location of failure. For instance, in the case of a trunk failure, all downstream and upstream channels experience loss. However, in the case of a distribution link failure, downstream channels destined to those ONUs that are located on the ring right after the failure point are the only ones that experience loss. On the contrary, upstream channels corresponding to those ONUs that are located on the ring before the failure point are the only ones that experience loss (Note that all downstream, upstream, and LAN signals are transmitted in clockwise direction). Thus, in the case of a last/first link failure, all upstream/downstream channels experience loss but none of the downstream/upstream channels does. On the other hand, in the case of a node failure, none of the down/upstream channels experience loss except for the downstream channel that is being terminated (dropped) at the failed node. This is due to the fact that nodes are uncoupled from the ring via OADM which is attached to each node (see figure 6.1).

Figure 6.5 shows total downstream traffic loss versus average down-link network load for several different network failure scenarios including trunk failure, first link failure, fifth link failure(link between ONU_4 and ONU_5), fifteenth link failure (link between ONU_{15} and ONU_{16}), and node failure. Note that downstream traffic loss depends on four factors: 1) offered network load; the higher the load the higher the traffic loss. 2) existing traffic on the trunk and distribution ring at the time of the failure. 3) the duration from when a failure occurs to when the OLT detects absence of the LAN signal (only in the case of the 3 special failures, trunk, last node, and last link) or receives a failure indication message from the affected ONU instructing it to stop further downstream transmissions (in all other failure cases). 4) type and location of failure.

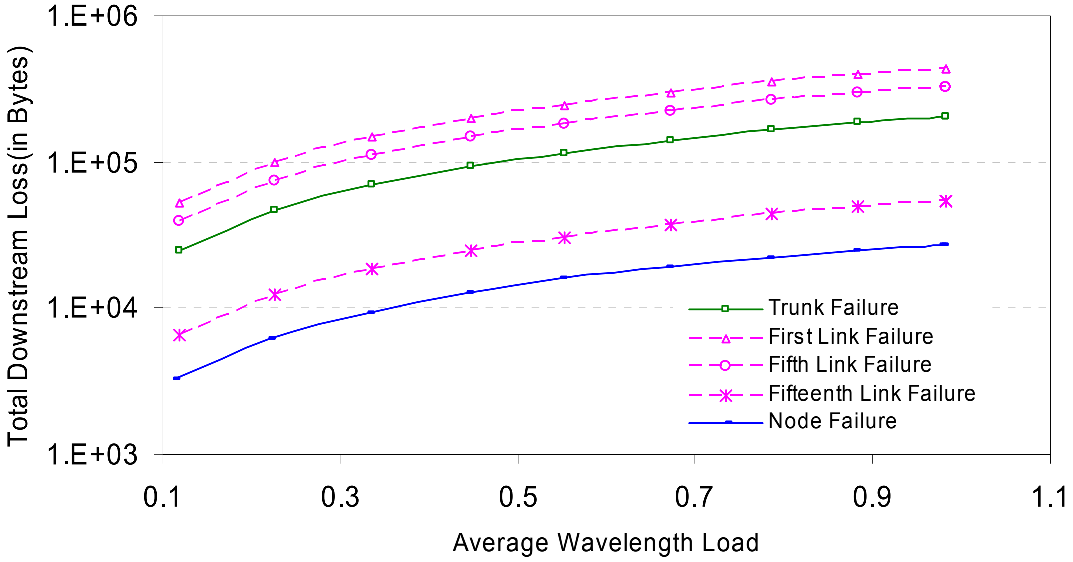


Figure 6.5. Average downstream traffic loss for different failure scenarios

As expected, as can be seen from Fig. 6.5, a node failure causes the least amount of traffic loss while a first link failure causes the most. Downstream traffic experiences the maximum loss in the case of the first link failure for the following two reasons: 1) all downstream channels experience loss; and 2) the duration of traffic loss, i. e., the interval from when a failure occurs to when the OLT receives a failure indication message from ONU_1 instructing it to stop further downstream transmissions is the longest (maximum propagation delay between $ONU_1$and the OLT). In this case, average traffic loss per channel includes existing traffic on the trunk fiber at the time of failure (~100 µs) as well as traffic transmitted by OLT during the duration form the instant of failure occurrence until it receives alarm indication message from ONU_1 instructing it to stop further downstream transmissions (5 µs + 115 µs + 0.5 µs ~ 121 µs). Thus, total downstream traffic loss per channel in the case of a first link failure is equivalent to ~221 µs of transmitted traffic by OLT.

On the other hand, in the case of a node failure, none of the downstream channels experience loss except for the downstream channel that is being terminated (dropped) at the failed node and, thus, downstream traffic loss is minimal. Analogous to the first link failure, though all downstream traffic experiences loss in the case of a trunk failure, however, as can be seen form Figure 6.5, the traffic loss is less than that of a first link failure. This is because the duration of traffic loss in the case of a trunk failure (~55 µs) is less than that of a first link failure (~ 121 µs). Assuming that fiber cut occurs at the middle of the trunk, downstream traffic loss per channel includes existing traffic already on the trunk fiber at the time of failure (~50 µs) as well as traffic transmitted by OLT during the duration form the instant of failure occurrence up to the time when it is able to detect the failure and stops further transmission (detection time + timeout ~55 μs). Thus, total downstream traffic loss per channel in the case of a trunk failure is equivalent to ~105 μs of transmitted traffic by OLT.

Figure 6.6 shows total upstream traffic loss versus average network load for several different network failure scenarios including trunk failure, last link failure, fifth link failure, fifteenth link failure and last link failure. Note that upstream traffic loss also depends on the same four factors listed above except that it is now the affected ONU that must detect the failure and notify all other ONUs to stop upstream transmissions. As can be seen from Figure 6.6, in contrast to downstream traffic loss, the fifth link failure causes the least amount of traffic loss while a trunk failure causes the most. This is because the average time it takes for ONU_1 to detect a trunk failure (50 μs) is much longer than that taken by any other ONU to detect any and all other general failure scenarios (~ 0 sec).

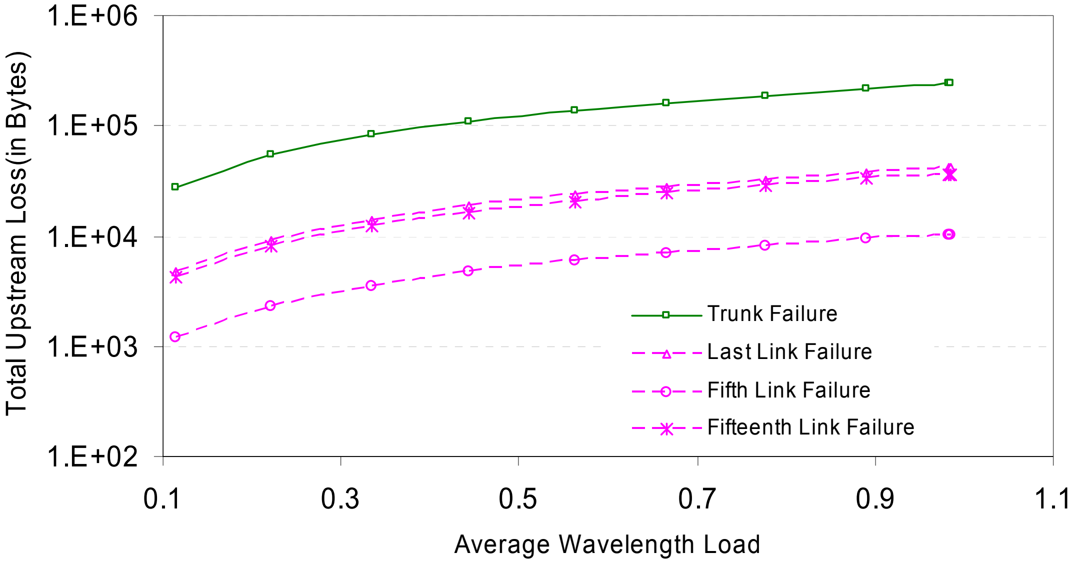


Figure 6.6. Average upstream traffic loss for different failure scenarios

Assuming that a fiber cut occurs at the middle of the trunk, average upstream traffic loss per channel includes existing traffic on the fibers (trunk + distribution ring) at the time of failure (equivalent to ~ 65 μs of transmitted traffic), traffic lost during ONU_1's detection time (equivalent to 55 μs of transmitted traffic), traffic lost during ONU_1's REPORT transmission time (equivalent to ~0.5 μs of transmitted traffic), traffic lost during REPORT propagation to the ONUs on the ring, plus average traffic lost during REPORT'S propagation time around the ring to notify all other ONUs to stop upstream transmission (equivalent to ~1 to 15 μs of transmitted traffic). Thus, average upstream traffic lost per ONU in the case of a trunk failure is equivalent to ~ 121 μs of transmitted traffic by the ONUs.

In the case of last link failure, upstream traffic lost includes existing traffic on the ring at the time of failure (equivalent to ~1-15 μs of transmitted traffic depending on whether the failure occurs near the beginning or the end of the ring, respectively), traffic lost due to timeout period (~5 μs), traffic lost during ONU_1's REPORT transmission time (equivalent to 0.5 μs of transmitted traffic), plus traffic lost during REPORT's propagation time around the ring to notify all other ONUs to stop upstream transmission (equivalent to ~1-15 μs of transmitted traffic). Thus, in the case of a general link/node failure, average upstream traffic loss per ONU is equivalent only to ~ 22 μs of transmitted traffic by ONUs.

Fig. 6.7 shows LAN traffic loss versus offered LAN load for a node/distribution link failure scenario. In the case of a node or distribution link failure, LAN traffic loss includes following components; 1) traffic loss due to existing traffic on the ring at the time of failure (~1 to 15 μs), 2) traffic loss due to timeout period (5 μs), 3) traffic loss due to REPORT transmission time (0.5 μs), 4) traffic loss due to REPORT propagation time to all other ONUs (~1 to 15 μs). Thus total LAN traffic loss is equivalent to ~21 μs of transmitted traffic by ONUs.

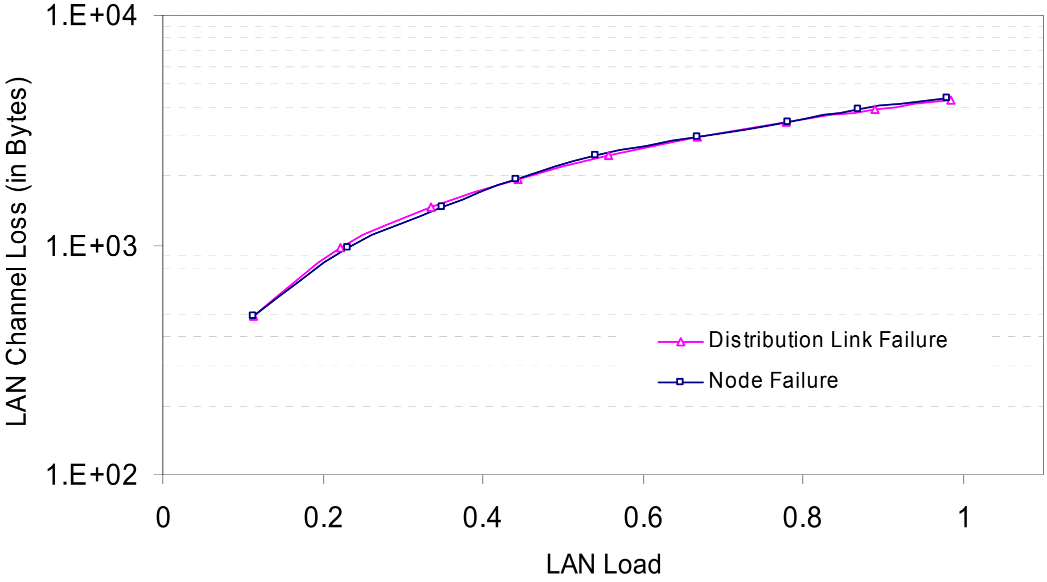


Figure 6.7. LAN traffic loss based on node and distribution link failure

6.7 Conclusion

This chapter proposed two different self healing architectures for the proposed WDM-PON ring architecture which are fully distributed self healing WDM-PON ring architecture and hybrid self healing WDM-PON ring architecture. Both architectures provide simple and cost-effective resilience capabilities against any and all kinds of networking failures. Both of the proposed schemes are capable of protecting against both node (ONU) and distribution/trunk fiber failures. These schemes enable the restoration of all network traffic including upstream, downstream, and LAN data. The proposed schemes can also protect against any combination of concurrent double failures including trunk/distribution fiber breaks and node failures.

The simulation results indicate that the performance of the both schemes, including traffic loss and restoration speed outperforms the standard schemes which are recommended by ITU-T G.983.1. Furthermore, the recovery time associated with any and all different distribution network/trunk failures is still within the delay-bound limit required for delivering guaranteed triple play services.

Chapter 7

Conclusions and Future Work

7.1 Conclusion

This dissertation has examined the technological requirements and assessed the performance analysis and feasibility for implementing a novel WDM-PON-based FTTH broadband access architecture that addresses the key limitations of conventional tree-based PON architectures including: a) providing a simple and cost-effective fully distributed/hybrid resilience capabilities against most types of networking failures; b) supporting private networking capability as well as dynamic allocation and sharing of network resources. Specifically, we have proposed and devised a novel self healing ring-based local access C/WDM-PON architecture that efficiently supports dynamic allocation of wavelengths/ timeslots and sharing traffic as well as truly direct private connections among PON end users. The proposed architecture combines the salient features of both traditional static WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and increased security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end users). The simulation results have indicated that the performance of the proposed WDM-PON ring scheme in terms of throughput can exceed the dedicated capacity by making use of unused capacity of idle or lightly loaded channels.

We have also proposed two different self healing architectures for the proposed WDM-PON ring architecture, namely fully distributed self healing WDM-PON ring architecture and hybrid self healing WDM-PON ring architecture. Both architectures provide simple and cost-effective resilience capabilities against any and all kinds of networking failures. Both of the proposed schemes are capable of protecting against both node (ONU) and distribution/trunk fiber failures. These schemes enable the restoration of all network traffic including upstream, downstream, and LAN data. The proposed schemes can also protect against any combination of concurrent double failures including trunk/distribution fiber breaks and node failures. The simulation results have indicated that the performance of the both schemes, including traffic loss and restoration speed outperforms the standard schemes which are recommended by ITU-T G.983.1. Furthermore, the recovery time associated with any and all different distribution network/trunk failures is still within the delay-bound limit required for delivering guaranteed triple play services.

7.2 Future Work

In this dissertation, we proposed the ring-based WDM-PON solution to some of the obstacles of legacy WDM-PON architectures. The theoretical analysis and simulation results were presented to demonstrate the advantages of the proposed approach. Although results of simulation experiments were shown, the ultimate test would require implementing the proposed algorithms in real network with realistic traffic. Therefore, physical experiments (test bed) to observe real life implications of the proposed approach would be an avenue of future investigation.

One of the future works would be to develop a quality of service (QoS) algorithm over proposed WDM-PON ring architecture. Providing triple play services to end-user by meeting the traffic engineering criteria of each class of service (CoS) is very important in today's networks

Another area of future work would be network security. The way the proposed architecture works is that one ONU traffic (LAN/WAN) passes through the other ONUs. It raises security concern. Therefore, one of the areas to explore in future would be security related issues on this architecture.

Abbreviations

ADSL Asymmetric Digital Subscriber Line

ATM Asynchronous Transfer Mode

AON Active Optical Network

APON Asynchronous Transfer Mode Optical Network

APS Automated Protection Switching

AWG Arrayed Waveguide Grating

BPON Broadband Passive Optical Network

BE Best Effort

CDMA Code Division Multiplexing Access

CMTS Cable Modem Termination System

CO Central Office

CRC Cyclic Redundancy Check

CWDM Coarse Wavelength Division Multiplexing

DBA Dynamic Bandwidth Allocation

DRR Deficit Round Robin

DSL Digital Subscriber Loop

DSLAM Digital Subscriber Line Access Multiplexer

DWDM Dense Wavelength Division Multiplexing

EDFA Erbium Doped Fiber Amplifier

EFMA Ethernet in the First Mile Alliance

EPON Ethernet Passive Optical Network

ESCON Enterprise System Connection

FDDI Fiber distributed data interface

FEC Forward Error Correction

FICON Fiber Connectivity

FIFO First In First Out

FN Fiber Node

FSAN Full Service Network Access

FTTB Fiber-to-the-Building

FTTC Fiber-to-the-Curb

FTTCab Fiber-to-the-Cabinet

FTTH Fiber-to-the-Home

FTTN Fiber-to-the-Node

GPON Gigabit Passive Optical Network

HFC Hybrid Fiber Coaxial

IFG Inter-frame Gap

IEEE Institute of Electric and Electronic Engineers

IETF Internet Engineering Task Force

IPACT Interleaved Polling with Adaptive Cycle Time

IPTV Internet Protocol Television

ISDN Integrated Services Digital Network

ITU International Telecommunications Union

LAN Local Area Network

LLC Logical Link Control

MAC Media Access Control

MAN Metropolitan Area Network

MPCP Multi-Point Control Protocol

NRZ Non Return to Zero

NSPQ Non Strict Priority Queuing

OADM Optical Add Drop Multiplexer

OAM Operations and Maintenance

OLT Optical Line Terminal

ONU Optical Network Unit

ONT Optical Network Terminal

OUT ONU Under Test

PON Passive Optical Network

POTS Plain Old Telephony System

PRBS Pseudo Random Bit Sequence

PtMP Point-to-Multi Point

QoS Quality of Service

SC Star Coupler

SDMA Subcarrier Division Multiplexing Access

SLA Service Level Agreement

SNR Signal to Noise Ratio

SP Service Provider

SPQ Strict Priority Queuing

SRF Shortest Request First

SWS Shared Wavelength Scheduling

RTT Round Trip Time

TC Transient Counter

TDMA Time Division Multiplexing Access

TDS Transient Downstream Traffic

TFF Thin Film Filter

TUS Transient Upstream Traffic

VDSL Very High Speed DSL

VoD Video on Demand

VOIP Voice over Internet Protocol

WAN Wide Area Network

WDM Wavelength Division Multiplexing

WDMA Wavelength Division Multiplexing Access

Publications

- **Hasan Erkan** et al, “On the Merits of Implementing a Novel Decentralized Ethernet-based PON Architecture for Next-Generation Broadband Access Networks,” SPIE Optics East 2005 Symposium
- Hossain, D.; **Erkan, H**.; Dorsinville, R.; Ali, M.; Shami, A. “A novel ring-based EPON architecture”, Broadband Networks, 2005 2nd International Conference on, IEEE.
- Hossain, D.; **Erkan, H**.; Dorsinville, R.; Ali, M.; Shami, S.; Assi, C., “Protection for a ring-based EPON architecture”, Broadband Networks, 2005 2nd International Conference on, IEEE.
- **Hasan Erkan** et al, “A simple and Cost Effective Ring-Based Local Access C/DWDM-PON Architecture for Supporting A Truly Shared LAN Capability” Military Conference (Milcom)2007, Orlando
- **Hasan Erkan** et al, “A Novel Ring-Based WDM-PON Access Architecture for the Efficient Utilization of Network Resources” International conference on Communication (ICC-2008), Beijing, China
- Hossain, D.; **Erkan, H.** et al “A Survivable Broadband Local Access PON Architecture: A New Direction for Supporting Simple and Efficient Resilience Capabilities”, Journal of Optical Networks (JON), (Submitted)
- **Erkan, H.** et al “A Novel Cost Effective Open Access CWDM-PON Architecture”, 2008 IEEE Sarnoff Symposium, Nassau Inn in Princeton, NJ, USA (Submitted)

Bibliography

Chapter 1

[1] B. Lung, “PON architecture ‘future proofs’ FTTH,” Lightwave, PennWell, 16(10):104 –107, 1999.

[2] G. Kramer and G. Pesavento, “Ethernet Passive Optical Network (EPON): Building a Next Generation Optical Access Network,” IEEE Com. Mag., pp. 66-73, Feb. 2002.

[3] G. Kramer et al., “Interleaved Polling with Adaptive Cycle Time (IPACT): A Dynamic Bandwidth Distribution Scheme in an Optical Access Network,” Photonic Network Comm, 4(1): 89-107, 2002

[4] C. Assi et al., “Dynamic bandwidth allocation for quality of service over Ethernet PONs,” IEEE J. Select. Areas Commun., 21(9):1467- 1477, 2003

[5] Chang-Hee Lee, et al., “Fiber to the home using a PON infrastructure,” J. Lightw. Technol., 24(12):4568-4583, 2006.

[6] Huan Song, et al., “Shared-wavelength WDM-PON access network for supporting downstream traffic with QoS” Proc. OFC/NFOEC, Anaheim, CA, March 2006.

[7] Huan Song, et al., "Shared Wavelength WDM-PON Access Network", Proc. COIN, Korea, 2006.

[8] M. McGarry, et al., An Evolutionary Upgrade for EPONs. Technical Report, Dept of Electrical Engineering, Arizona State University, Sept. 2004.

[9] M. McGarry, et al., “WDM Ethernet Passive Optical Networks,” IEEE Com. Mag., pp. 518-525, Feb. 2006.

[10] Dhaini, et al., “Dynamic Wavelength and Bandwidth Allocation in Hybrid TDM/WDM EPON Networks”, J. Lightw. Technol., 25(1):277 – 286, 2007.

[11] C.-J. Chae et al., “A PON system suitable for internetworking optical network units using a fiber Bragg grating on the feeder fiber,” IEEE Photon. Technol. Lett., 11(12):1686–1688, 1999.

[12] D. sala, et. Al., IEEE EFM, May 2002 materials, www. ieee802.org/3/efm.

[13] C.-J. Chae et al., “A PON system suitable for internetworking optical network units using a fiber Bragg grating on the feeder fiber,” IEEE Photon. Technol. Lett., vol. 11, pp. 1686–1688, Dec. 1999.

[14] E. Wong et al., “CSMA/CD-based EPON with optical internetworking capability among users,” IEEE PTL, Vol 16, No 9, pp. 2195-97, Sep. 2004.

[15] Qiguang Zhao et al., “A novel internetworking scheme for WDM passive optical network based on remodulation technique” Proc. OFC/NFOEC, Anaheim, CA, March 2006 Page(s):3 pp.

[16] Journal of Selected Area of Communication, 2007

[17] T. J. Chan, C. K. Chan, L. K. Chen and F. Tong, “A self-protected architecture for wavelength-division-multiplexed passive optical networks,” IEEE Photon. Tech. Lett., vol. 15, pp. 1660-1662, Nov. 2003.

[18]Z. X. Wang, X. F. Sun, Chinlon Lin, C. K. Chan, L. K. Chen, “A novel centrally controlled protection scheme for traffic restoration in WDM passive optical networks,” IEEE Photon. Tech. Lett., vol. 17, no. 3, pp. 717-719, Mar. 2005.

[19] C. M. Lee, T. J. Chan, C. K. Chan, L. K. Chen, and C. Lin, “A group protection architecture (GPA) for traffic restoration in multi-wavelength passive optical network,” in ECOC 2003, Rimini, Italy, 2003, Paper Th2.4.2.

[20] E. S. Son, K. H. Han, J. H. Lee, and Y. C. Chung, “Survivable network architectures for WDM PON”, OFC/NFOEC’05, Paper OFI4, Anaheim, California, USA, 2005.

[21] X. F. Sun, C. K. Chan, L. K. Chen, “A survivable WDM PON architecture with centralized alternate-path protection switching for traffic restoration,” IEEE Photon. Tech. Lett., vol. 18, no. 4, pp.631-633, Feb. 2006.

[22] Chun-Kit Chan, Lian-Kuan Chen, and Chinlon Lin “WDM PON for Next-Generation Optical Broadband Access Networks,” The Eleventh Optoelectronics and Commun. Conference (OECC 2006), 5E2-1-1, July 3-7, 2006, Kaohsiung, Taiwan.

Chapter 2

[1] X. Lin and O. Sneizko, “The evolution of cable TV networks,” in Optical Fiber Communications IVB, I. Kaminow and T. Li, Eds. New York: Academic, 2002

[2] K. G. Coffman and A. M. Odlyzko, “Internet growth: Is there a "Moore's Law" for data traffic?” Handbook of Massive Data Sets, J. Abello, P. M. Pardalos, and M. G. C. Resende, eds., Kluwer, 2001.

[3] “Access Network Systems: North America – Optical Access. DLC and PON Technology and Market Report,” Report RHK-RPT-0548, RHK Telecommunication Industry Analysis, San Francisco, June 2001.

[4] G. Pesavento and M. Kelsey, “PONs for the broadband local loop,” Lightwave, PennWell, vol. 16, no. 10, pp. 68 – 74, September 1999.

[5] B. Lung, “PON architecture ‘futureproofs’ FTTH,” Lightwave, PennWell, vol. 16, no. 10, pp. 104 –107, September 1999.

[6] IEEE 802.3 Ethernet in the First Mile Study Group [Online]. Available: http://www.ieee802.org/3/efm/public/index.html.

[7] J.R. Stern, J.W. Ballance, D.W. Faulkner, S. Hornung, and D.B. Payne, “Passive Optical Local Networks for Telephony Applications and Beyond,” Electronics Letters, vol. 23, no. 24, pp. 1255–1257, Nov. 1987.

[8] Donald. E.A. Clarke and Tetsuya Kanada, “Broadband: The Last Mile,” IEEE Communications Magazine, vol. 31, no. 3, pp. 94–100, Mar. 1993.

[9] Yih-Kang Maurice Lin, Dan R. Spears, and Mih Yin, “Fiber-Based Local Access Network Architectures,” IEEE Communications Magazine, vol. 27, no. 10, pp.64–73, Oct. 1989.

[10] Ivan Andonovic and Deepak Uttamchandani, Eds., Principles Of Modern Optical Systems, chapter 11, Artech House, 1989.

[11] G. Kramer and G. Pesavento, “Ethernet passive optical network (EPON): building a next generation optical access network,” IEEE Commun. Mag., pp. 66–73, Feb. 2002.

[12] S. Clavenna, “Metro Optical Ethernet,” Lightreading (www.lightreading.com), November 2000.

[13] http://lw.pennnet.com

[14] E. Harstead and P. H. van Heyningen, “Optical access network,” in Optical Fiber Communications IVB, I. Kaminow and T. Li, Eds. New York: Academic, 2002.

Chapter 3

[1] Full Services Access Networks, http://www.fsanet.net/.

[2]IEEE 802.3 Ethernet in the First Mile Study Group, http://www.ieee802.org/3/efm/public/index.html

[3] B. Lung, “PON Architecture ‘Future proofs’ FTTH,” Lightwave, vol. 16, no. 10, Sept. 1999, pp. 104-7.

[4] D. Mynbaev, “From Core to Metro to Access Networks – The Need for Passive Optical Networks”.

[5] G. Kramer and G. Pesavento, “Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Access Network,” IEEE Com. Mag., pp. 66-73, Feb. 2002.

[6] Alloptic, “Ethernet Passive Optical Networks,” The International engineering Consortium, http://www.iec.org .

[7] G. Kramer, et al. “Ethernet PON: Design and Analysis of an Optical Access Network,” Photonic Network Communications Journal, vol. 3, No. 3, July 2001

[8] G. Kramer, B. Mukherjee, and G. Pesavento, “IPACT: A dynamic Protocol for an Ethernet PON (EPON),” IEEE Comm. Mag., pp. 74-80, Feb. 2002.

[9] A. Hadjiantonis, T. Rahman, A. Khalil, G. Ellinas, M. F. Arend and M. A. Ali, “A Novel Decentralized Ethernet-Based PON Architecture” Proc. IEEE ICC’04, Paris, France, June 2004.

[10] Chae Chang-Joon, Elaine Wong, Rodney S. Tucher, “Optical CSMA/CD Media Access Scheme for Ethernet Over Passive Optical Network,” IEEE Photonics Technology Letters, vol. 14, No.5, Mat 2002.

[11] D. Marcuse, “Derivation of analytical expressions for bit error probability in lightwave systems with optical amplifier” , Journal of Lightwave Technology , Vol.8, Dec 1990.

Chapter 4

[1] P. J. Winzer, F. Fidler, M.J. Matthews, L. E. Nelson, S. Chandrasekhar, L. L. Buhl, M. Winter, D. Castagnozzi, "Electronic equalization and FEC enable bidirectional CWDM capacities of 9.6 Tb/s-km " OFC 2004, Volume 2, 23-27 Feb. 2004 Page(s):3 pp. vol.2

[2]Amitabha Banerjee, Biswanath Mukherjee, Glen Kramer "Achieving Open Acces in Ethernet PON (EPON)" OFC 2005

[3] SFP PROTM Chipset and Reference Design, (Vitesse Semiconductor Corporation, 2004), ttp://www.vitesse.com/technologies/files/8/SFPPROChipsetandRefDesign.pdf.

[4] M. Shreedhar and G. Varghese, "Efficient Fair Queuing using Round Robin," Proc., ACM SIGCOMM '95, Cambridge, MA, USA.

[5] A. Demers, S. Keshav, and S. Shenker. "Analysis and simulation of a fair queuing algorithm. ACM SIGCOMM´89, pp 3-12.

[6] Arunabha Sen, Ibraz Mohammed, Ravikanth Samprathi, Subir Bandyopadhyay. "Fair Queuing with Round Robin: A New Packet Scheduling Algorithm for Routers, Proceedings of the Seventh International Symposium on Computers and Communications (ISCC'02)

[7] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: A dynamic Protocol for an Ethernet PON (EPON)," IEEE Comm. Mag., pp. 74-80, Feb. 2002.

Chapter 5

[1] B. Lung, “PON architecture ‘future proofs’ FTTH,” Lightwave, PennWell, 16(10):104 –107, 1999.

[2] G. Kramer and G. Pesavento, “Ethernet Passive Optical Network (EPON): Building a Next Generation Optical Access Network,” IEEE Com. Mag., pp. 66-73, Feb. 2002.

[3] G. Kramer et al., “Interleaved Polling with Adaptive Cycle Time (IPACT): A Dynamic Bandwidth Distribution Scheme in an Optical Access Network,” Photonic Network Comm, 4(1): 89-107, 2002

[4] C. Assi et al., “Dynamic bandwidth allocation for quality of service over Ethernet PONs,” IEEE J. Select. Areas Commun., 21(9):1467- 1477, 2003

[5] Chang-Hee Lee, et al., “Fiber to the home using a PON infrastructure,” J. Lightw. Technol., 24(12):4568-4583, 2006. [1] B. Lung, “PON architecture ‘future proofs’ FTTH,” Lightwave, PennWell, vol. 16, no. 10, pp. 104 –107, September 1999.

[6] Huan Song, et al., “Shared-wavelength WDM-PON access network for supporting downstream traffic with QoS” Proc. OFC/NFOEC, Anaheim, CA, March 2006.

[7] Huan Song, et al., "Shared Wavelength WDM-PON Access Network", Proc. COIN, Korea, 2006.

[8] M. McGarry, et al., An Evolutionary Upgrade for EPONs. Technical Report, Dept of Electrical Engineering, Arizona State University, Sept. 2004.

[9] M. McGarry, et al., “WDM Ethernet Passive Optical Networks,” IEEE Com. Mag., pp. 518-525, Feb. 2006.

[10] Dhaini, et al., “Dynamic Wavelength and Bandwidth Allocation in Hybrid TDM/WDM EPON Networks”, J. Lightw. Technol., 25(1):277 – 286, 2007.

[11] C.-J. Chae et al., “A PON system suitable for internetworking optical network units using a fiber Bragg grating on the feeder fiber,” IEEE Photon. Technol. Lett., 11(12):1686–1688, 1999.

[12] E. Wong et al., “CSMA/CD-based EPON with optical internetworking capability among users,” IEEE Photon. Technol. Lett., 16(9):2195-97, 2004.

[13] N. Nadarajah, et al., “Novel schemes for local area network emulation in passive optical networks with RF subcarrier multiplexed customer traffic,” J. Lightw. Technol., 23(10):2974–2983, 2005.

[14] A. Delowar, et al., “Supporting Private Networking Capability in EPON”, Proc. ICC,Istanbul,Turkey,June2006.

[15] Qiguang Zhao et al., “A novel internetworking scheme for WDM passive optical network based on remodulation technique” Proc. OFC/NFOEC, Anaheim, CA, March 2006 Page(s):3 pp.

[16] Journal of Selected Area of Communication, 2007

[17] G. Pesavento and M. Kelsey, “PONs for the broadband local loop,” Lightwave, PennWell, vol. 16, no. 10, pp. 68 – 74, September 1999.

[18] A. Delowar, et al., “Ring-based local access PON architecture for supporting private networking capability”, Journal of Optical Networking, Vol. 5, Issue 1, pp. 26-3

Chapter 6

[1] D. Sala and A. Gummalla, “PON functional requirements: services and performance,” in IEEE 802.3ah Meeting in Portland OR, July 2001.

[2] G. Kramer, B. Mukherjee, and G. Pesavento, “IPACT: A dynamic Protocol for an Ethernet PON (EPON),” IEEE Comm. Mag., pp. 74-80, Feb. 2002.

[3] G. Kramer et al., “Ethernet PON: design and analysis of an optical access network,” Photon. Network Commun. J., vol. 3, no. 3, July 2001.

[4] C. Assi et al., “Dynamic bandwidth allocation for quality of service over Ethernet PONs,” IEEE J. Select. Areas Commun., Dec. 2003.

[5] M. K. Abdullah, W. T. P’ng, P. W. Lau, and E. R. Tee, “FTTH access network protection using a switch,” in Proc. 9th Asia-Pacific Conf. Commun., 2003, pp. 1219–1222.

[6] W.-P. Lin, M.-S. Kao, and S. Chi, “The modified star-ring architecture for high-capacity subcarrier multiplexed passive optical networks,” J. Lightw. Technol., vol. 19, no. 1, pp. 32–39, Jan. 2001.

[7] W.-P. Lin, M.-S. Kao, and S. Chi, “A DWDM/SCM self-healing architecture for broad-band subscriber networks,” J. Lightw. Technol., vol. 21, no. 2, pp. 319–328, Feb. 2003.

[8] ITU-T recommendation G.983.1, "Broadband optical access systems based on passive optical networks (PON): 1998.

[9] Yu-mi Kim et al., “Cost Effective Protection Architecture to Provide Diverse Protection Demands in Ethernet Passive Optical Network”, Proceedings of ICCT 2003.

[10] N. Nadarajah, M. Attygalle, E. Wong, and A. Nirmalathas, “Novel schemes for local area network emulation in passive optical networks with RF subcarrier multiplexed customer traffic,” J. Lightw. Technol., vol. 23, no. 10, pp. 2974–2983, Oct. 2005.

[11] O. Gerstel and R. Ramaswami, “Optical layer survivability: A post-bubble perspective,” IEEE Commun. Mag., vol. 41, no. 9, pp. 51–53, Sep. 2003.

[12] T.-H. Wu, “Emerging technologies for fiber network survivability,” IEEE Commun. Mag., vol. 33, no. 2, pp. 62–74, Feb. 1995.

[13] Alliance Fiber Optic, http://www.afop.com, AC Photonics, http://www.acphotonics.com, http://www.dbmoptics.com/products/gpoti/switches.php

[14] A. Delowar Hossain, A. Shami, C. Assi, and M. A. Ali "Supporting Private Networking Capability in EPON", *Proc. International Communications Conference (ICC'2006)*, Istanbul, Turkey, June 2006.