Wavelength deployment schemes in next generation PONs

March 2023

ZHANG CHENXIAO

Wavelength deployment schemes in next generation PONs

Graduate School of Science and Technology
Degree Programs in Systems and Information Engineering
University of Tsukuba

March 2023

ZHANG CHENXIAO

Abstract

With the development of Internet services such as machine-to-machine communication, ultra-high-definition video, on-line games and so on, the global Internet traffic has been experiencing a rapid growth. From the concept being proposed in the late 1980s, passive optical networks (PONs) have been developing to address the continuous growth of network traffic. Advances in optical components, modulation formats and other technologies have enabled transceivers in commercial PONs with bit rate up to 10 Gbps. However, increasing the bit rate over 10 Gbps is limited by the current economic ecology, that is, too expensive for commercial PONs. Therefore, wavelength division multiplexing technology is introduced in next generation PONs to increase the total capacity, which brings a significantly evolution to the architecture of PONs. In addition, PONs have the potential to support mobile x-haul transmission in a cost-effective manner. However, mobile x-haul transmission, especially the fronthaul transmission, has stringent latency and bandwidth requirement, which brings new challenges for next generation PONs.

In this thesis, we mainly focus on the design problem of next generation PONs, that is, the resource assignment problem in the static scenario. For PONs, resource assignment problem occurs only in upstream transmission due to the multiple access technology used in upstream transmission. We define problems based on the characteristic of the PON architectures and the traffic requests.

In the first research, we focus on the trade-off between ONU cost and the upstream performance in a static scenario, and we develop an integer linear programming (ILP) model to formulate the problem. In addition, we propose a heuristic algorithm to effectively select appropriate wavelength deployment schemes. From our simulation results, we can see that our proposed algorithm can efficiently find a finer number of deployment wavelengths to balance the cost of ONUs and the upstream transmission performance. The wavelength deployment schemes for future PONs can also be specified based on whether network operators prioritize network performance or cost.

In the second research, we concentrate on how to design fixed-mobile convergence (FMC) access networks based on two potential architectures proposed in previous works and aim at minimizing the number of wavelengths required for the FMC access network. We develop an ILP model and obtain the lower bound of the solutions for this problem. Moreover, we propose a heuristic algorithm to calculate the results in large-scale networks. We numerically analyze the solutions with various network parameters, and the results show that the algorithm can effectively obtain the minimal number of required wavelengths while satisfying both fixed and mobile transmission latency requirements. Moreover, selecting an appropriate architecture leads to substantial cost savings for network operators.

Contents

List of Figures

List of Acronyms

ATM	**Asynchronous transfer mode**
BBU	**Baseband unit**
BPON	**Broadband PON**
BS	**Base station**
CO	**Central office**
CO-DBA	**Cooperative DBA**
C-RAN	**Cloud/centralized radio access network**
CU	**Central unit**
DBA	**Dynamic bandwidth assignment**
D-RAN	**Distributed radio access network**
DU	**Distributed unit**
DWBA	**Dynamic wavelength and bandwidth assignment**
EPON	**Ethernet PON**
FDD	**Frequency-division duplex**
FEC	**Forward error correction**
FF	**First-fit**
FMC	**Fixed-mobile convergence**
FSAN	**Full Service Access Network**
FTTB	**Fibre to the building**
FTTH	**Fibre to the home**
FW	**Flexible wavelength**
GEM	**GPON encapsulation mod**
GP	**Grant readjustment**
GPON	**Gigabit-capable PON**
HLS	**Higher-layer level split**
IEEE	**Institute of Electrical and Electronics Engineers**
ILP	**Integer linear programming**
IPACT	**Interleaved polling with adaptive cycle time**
ITU-T	**International Telecommunication Union Telecommunication Standardization Sector**
LLS	**Lower-layer split**
MAC	**Media access control**
MFH	**Mobile fronthaul**
MIMO	**Multiple-input multiple-output**
MNO	**Mobile network operator**

NC	**New core**
NFV	**Network functions virtualization**
NFVI	**NFV infrastructure**
NG-EPON	**Next generation EPON**
NG-PON2	**Next generation PON version 2**
ODN	**Optical distribution network**
OLT	**Optical line terminal**
ONU	**Optical network unit**
PON	**Passive optical network**
P2MP	**Point-to-multipoint**
P2P	**Point to point**
QoS	**Quality of service**
RAN	**Radio access network**
RRH	**Remote radio head**
RRU	**Remote radio unit**
RU	**Radio unit**
SCAP	**Single channel as possible**
SDN	**Software-defined networking**
SNR	**Signal to noise ratio**
TDD	**Time-division duplex**
TDM	**Time-division multiplexed**
TWDM	**Time- and wavelengh- division multiplexed**
VPON	**Virtual PON**
WDM	**Wavelength division multiplexing**
WDSCS	**Wavelength deployment scheme comparison and selection algorith**
WF	**Water-filling**
XG-PON	**10 gigabit PON**
XG-PON1	**10 gigabit PON version 1**
XG-PON2	**10 gigabit PON version 2**
XGS-PON2	**10 gigabit symmetrical PON**
5G	**Fifth-generation**

Chapter 1

Introduction

Passive optical network (PON) are massively deployed in many countries and regions to provide access services in a cost-effective manner [1]. Access network based on PONs can share trunk fibres and other passive elements so that to achieve low power consumption, deployment cost and so on. Most of the deployed commercial PONs adopt the time division multiplexing (TDM) technology to provide multiple access for upstream transmission. In addition, driven by the continuous growth in demand for capacity [2], next generation PONs will adopt wavelength division multiplexing (WDM) technology to extend the bandwidth over several wavelengths, which can significantly increase the total capacity of PONs. However, the expanded wavelength domain leads to many new challenges such as the problem of how to assign wavelength for upstream transmission. Another application of next generation PONs is to provide access services for mobile networks. However, mobile x-haul transmission (i.e., mobile fronthaul, midhaul or backhaul transmission) requires extremely low latency and very huge bandwidth, which leads to further changes in the resource assignment problem of the PONs.

In this chapter, we begin with an introduction to the field and the position of our research. Then, we briefly describe the main works in this thesis as well as the contribution of this thesis. Finally, we show the organization of this thesis.

1.1 Position of our research

Generally speaking, the communication network includes wireless network and wired network. The wireless network adopts radio waves as transmission media while wired networks use cables to transmit data. The wired transmission network can be further divided into the two categories: backbone network and access network. The backbone network provides connection between multiple areas or cities and the access network offers connection for subscribers to their network service provider. The PON is a network architecture of access network that can provide access service for subscribers in a cost-effective manner. Moreover, the research of PONs covers two fields: the physical layer as well as the network layer. Research on the former one includes the development of optics, modulation formats and so on; research on the latter one pays attention to the development of algorithms, the analysis of performance and so on. Our research is based on the network layer of PONs.

1.2 Main contributions of this thesis

In this thesis, we mainly focus on the design problems for next generation PONs. Specifically, we evaluate the device cost (influenced by the transmission technologies, modulation formats, multiplexing approach and so on) and the network performance (capacity, latency, security, extensibility, flexibility and other performance indicators) of PONs with different architectures in various situations, so that network operators can choose the most appropriate architecture to build the next-generation PONs based on their preference.

In the first work entitled "Analysis of wavelength deployment schemes in terms of ONUs cost and upstream transmission performance in NG-EPONs", we consider the trade-off between the network performance (i,e,. the upstream latency) and the device cost (i,e,. the ONUs cost). In addition, we provide a solution for network operators to select an appropriate architecture for next-generation PONs based on user requirements and quality of service (QoS) requirements.

In the second work entitled "Wavelength deployment scheme in a fixed-mobile convergence access network", we aim at minimizing the number of wavelengths required for the network while satisfying the latency requirement of both mobile x-haul transmission and fixed transmission. Moreover, we perform a comparison between the two potential architecture for the network, which can be used for as a reference when deploy the future PONs.

1.3 Organization of this thesis

The rest of the thesis is organized as follows. In chapter 2, we introduce the background of this thesis, including the technologies adopted in PONs and architectures of PONs. In chapters 3 and 4, we present the two works in the order of problem definition, previous works, problem formulation, proposed algorithm, numerical analysis and conclusion. In chapter 5, we summary our works in the research field. In chapter 6, we discuss the future work.

Chapter 2

Background

In this chapter, we first make an overview of the PONs. Then, we introduce the major innovations for next generation PONs, as well as the challenges carried by these innovative technologies. After that, we introduce the trend of PON applications in mobile x-haul transmission, as well as the issues brought by the trend.

2.1 Overview of the PONs

From the 1980s, optical transmission started being widely in the backbone networks to address the increasing traffic [3]. However, the capacity of the access network at that time lagged considerably behind the growth of Internet traffic, resulting in the "last mile" became the bottleneck of the whole network. The "last mile", which is now renamed to "first mile" by the Ethernet community to emphasize its importance, refers to the first part of connection from the subscriber side to the local equipment at the network operator side. In this context, the access network needed a new technology to provide the high-speed, high-capacity and highly flexible access services for subscribers in a cost-effective manner. PONs that enable the deployment of passive optical elements and share the trunk fibres connecting the centreal office (CO) and the user side can provide a cost-effective solution to bring fibre-to-the-building (FTTB) and fibre-to-the-home (FTTH) [4, 5].

2.1.1 Basic idea of the PONs

In a typical PON architecture based on tree topology, as shown in Fig. 2.1, an optical line terminal (OLT) in the CO and optical network units (ONUs) on the user side are connected via the optical distribution network (ODN). The elements in the ODN include trunk fibres and an 1: N optical combiner/splitter, and both of them are passive. In addition, PONs can be deployed in other topologies such as a ring, a bus

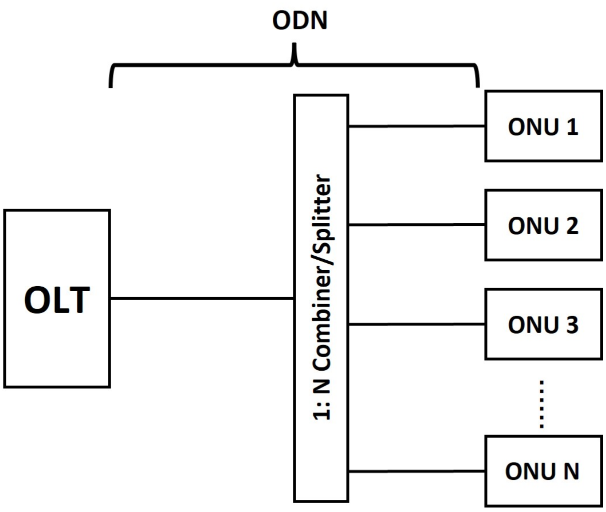

Figure 2.1: Typical architecture of a PON based on tree topology.

and so on.

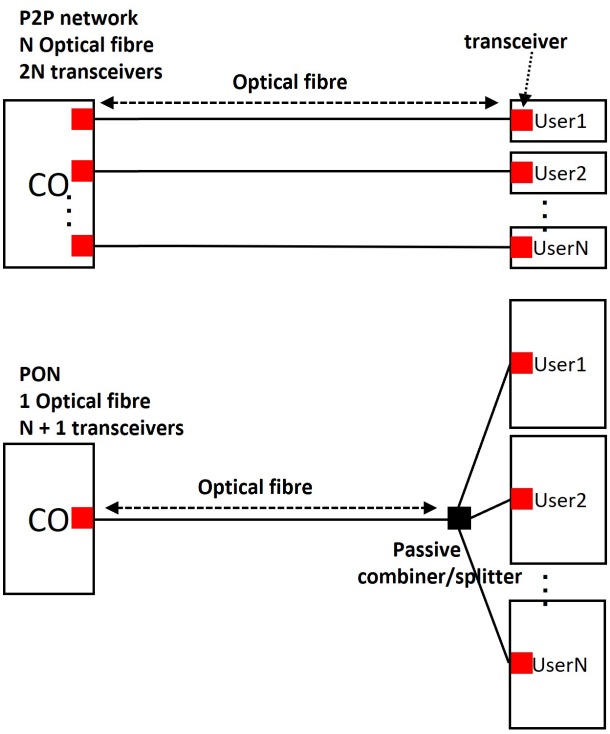

Figure 2.2: Implementation of FTTH via a PON or a point-to-point (P2P) network.

Fig. 2.2 shows the implementation of FTTH via a PON or a point-to-point (P2P) network. We can see that the PON network can provide significant savings in transceiver costs and fibre deployment costs compared to the P2P network. $2 \times N$ transceivers and N trunk fibre are required in the P2P network while N + 1 transceivers and 1 trunk fibre are required in the PON network. As a result, PONs can make great deployment cost savings mainly due to the sharing of the ODN element and transceivers[6].

2.1.2 Development of the PONs

The concept of the PON was first proposed in the late 1980s [7]. In 1997, the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) and the Full Service Access Network (FSAN) group [8] developed the first standard (G.983) of PONs [9]. The first version of G.893 used asynchronous transfer mode (ATM) as the signaling and transference protocol; therefore, it was also known as APON. Subsequently, the ITU-T/FSAN further amended and improved G.893 in 2001, and the final version of this standard was called broadband PON (BPON). The APON/BPON achieved the bit rate of 155.52 Mb/s for upstream and the bit rate of 155.52 Mb/s and 622.08 Mb/s for downstream. In 2003, the ITU-T/FSAN standardized the gigabit-class PON that was referred to as Gigabit-capable PON (GPON) [10]. The standard of GPON (G.984) adopted the GPON encapsulation mod (GEM) and offered 1.244 Gbps for upstream and 2.488 Gbps for downstream. Due to the development of Ethernet technology, the Institute of Electrical and Electronics Engineers (IEEE) established the " Ethernet in the First Mile" study group to developed the PON that was based on Ethernet technology. In 2004, the IEEE published 802.3ah and standardized Ethernet PON (EPON) [11], which had the bit rate of 1.25 Gbps for both upstream and downstream.

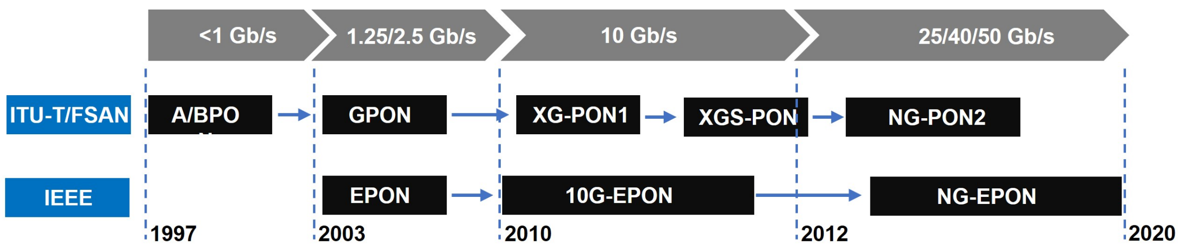

Figure 2.3: The development of PONs in the past few decades.

To satisfy the rapidly increasing amount of traffic, the ITU-T/FSAN and the IEEE standardized the 10 gigabit-class PON in the late 2000s. In 2010, the ITU-T/FSAN developed the G.987 that was called XG-PON [12]. XG-PON included an asymmetric version (XG-PON1) and a symmetric version (XG-PON2). XG-PON1 provided 2.488 Gbps for upstream and 9.953 Gbps for downstream; XG-PON2 provided 9.953 Gbps for both upstream and downstream but was not standardized. In 2017, the ITU-T/FSAN standardized the symmetric version for XG-PON (G.9807) which was known as XGS-PON ("S" means symmetric) [13]. In 2009, the IEEE published the 802.3av to standardize 10G EPON [14]. 10G EPON supported both the symmetric mode that provided 10 Gbps for both upstream and downstream and the asymmetric mode that provided 1 Gbps for upstream and 10 Gbps for downstream.

To further increase the capacity, the ITU-T/FSAN started the standardization of 40 gigabit-class PON in 2012. The 40G PON, which was also called NG-PON2, used the time- and wavelength- division multiplexing (TWDM) technology stacking 10 Gbps wavelengths to increase the peak bit rate to 10/40 Gbps for upstream and 40 Gbps for downstream. The standard was completed in 2015 (G.989) [15]. The IEEE established the 100G EPON Task Force (now it is known as 50G EPON Task Force) [16] in 2015 and aimed to standardize the next generation EPON (NG-EPON) to offered a cost-effective and practical solution for future EPONs. The objectives of NG-EPON included increasing the bit rate of a single channel to 25 Gbps without obvious increasing the cost per bit compared to existing 10G PONs as well as bonding 25 Gbps wavelengths to achieve aggregated peak bit rates of 25 Gbps and 50 Gbps for both upstream and downstream. In 2020, the standardization of NG-EPON was completed [17].

2.1.3 Data transmissions in the PONs

The downstream transmission from the OLT to ONUs is based on the point-to-multipoint (P2MP) manner [18], as shown in Fig. 2.4. In this case, the downstream transmission demands are encapsulated in frames (GEM frames in GPONs and Ethernet Frames in EPONs), which consists of header, data and trailer. The OLT broadcasts frames to ONUs pass through the 1: N passive optical combiner/splitter. Each ONU extracts the frames transmitted to it via the media access control (MAC) address information in the header of the frame.

The upstream transmission from ONUs to the OLT is similar to the P2P network. However, the collision may occur in the 1: N passive optical combiner/splitter when frames from different ONUs arrived simultaneously. Thus, multiple access approaches must be adopted in the PON, allowing ONUs to share all trunk fibre resources in the upstream transmission. Time-division multiplexed (TDM) technology is the most common approach to achieve multiple access in the PON [19], and the PON using TDM technology is called TDM-PON. Fig. 2.5 illustrates the upstream transmission in a TDM-PON, and frames from different ONUs are assigned to non-overlapping times slots to avoid the collision. Therefore, the resource problem of a PON occurs in the upstream transmission.

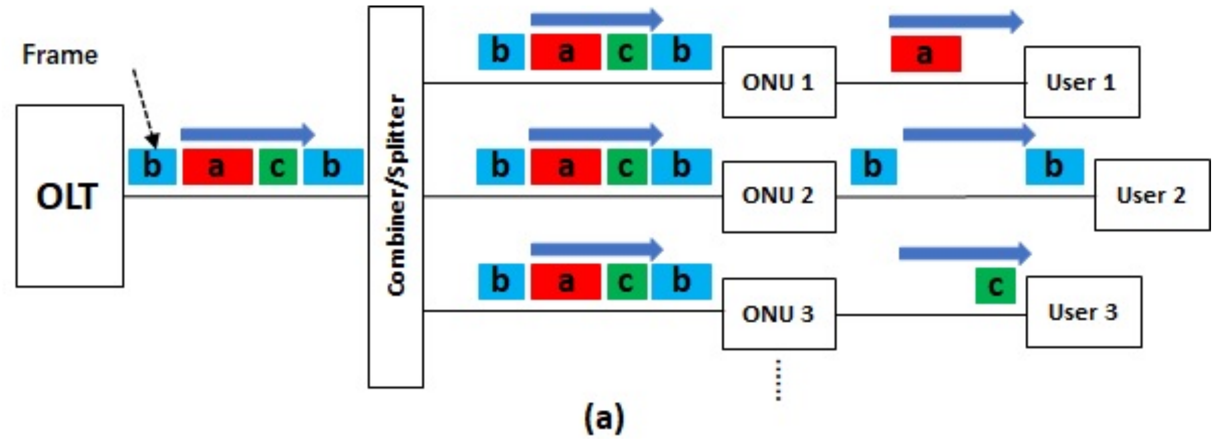

Figure 2.4: Downstream transmission in the PON.

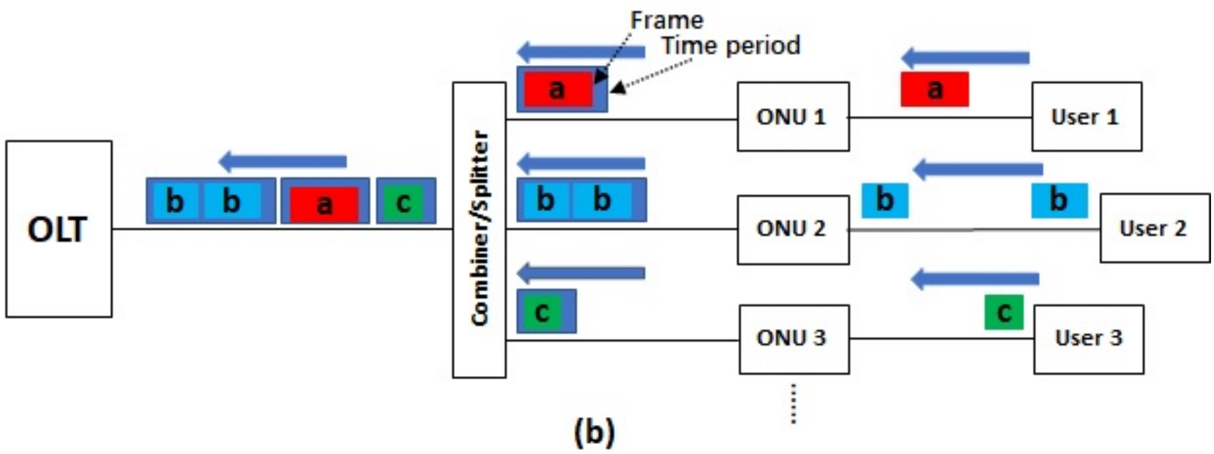

Figure 2.5: Upstream transmission in the PON.

2.1.4 Dynamic bandwidth assignment scheme

In the TDM-PON, the assignment of bandwidth to ONUs for upstream transmission is based on the dynamic bandwidth assignment (DBA) schemes adopted in the OLT [20]. According to the DBA scheme, the OLT dynamically assigned bandwidth to ONUs through the GATE and REPORT messages in a round-robin manner. In each cycle, ONUs send the REPORT messages, which consist of the information of the buffer state and the required bandwidth, to the OLT. When the OLT receives the REPORT messages from ONUs, the OLT assigns bandwidth to ONUs according to their requirement and sends the GATE messages to ONUs. The GATE message consists of the information about the time slots assigned to ONUs, and ONUs immediately transmit the frames to the OLT once ONUs receive the GATE message from the OLT. Moreover, a guard time (i.e., 1s) is included between two adjacent frames to avoid the overlap of the packets [21].

2.2 Next generation PONs

Over the past two decades, ODNs infrastructure for PONs has been widely deployed in many countries and regions, including China, Japan, Europe and the United States [22]. EPONs are mainly deployed in the United States, while GPONs are widely deployed in other countries and regions. Nowadays, most of the deployed commercial PONs are Gigabit-class PONs and being progressively updated to 10 Gigabit-class. However, driven by the rapid growth of the traffic demand, the ITU-T/FSAN and the IEEE have both investigated the next generation PONs.

2.2.1 Next generation PON2

In 2012, the ITU-T/FSAN has identified TWDM technology as the preferred solution for NG-PON2, and the PONs adopt the TWDM technology are called TWDM-PONs. As shown in Fig. 2.6, four 10 Gbps wavelengths (λ_1, λ_2, λ_3 and λ_4) are stacked to achieve the peak bit rake of 4×10 Gbps in a TWDM-PON. Stacking more than four wavelengths to provide higher capacity can also be achieved in NG-PON2. Besides the increased bandwidth capacity, TWDM-PONs provides higher flexibility compared to the deployed GPONs. For instance, TWDM-PONs support multiple services and user groups to be overlaid on the same fiber [23].

Wavelength turning is one of the technical enablers required for TWDM-PONs. To simplify operations, transceivers deployed in all ONUs must be colorless in a TWDM-PON [24]. That is, wavelength tunable rather than wavelength fixed transceivers are required for TWDM-PONs. In this way, the transceivers are able to tune to the correct wavelengths when transmitting frames in both upstream and downstream.

2.2.2 Next generation EPON

In 2020, the IEEE completed the standardization of NG-EPON with the two main objectives: increasing the bit rate of a single wavelength to 25 Gbps without a higher cost per bit than the current 10G PONs and multiplexing several 25 Gbps wavelengths to provide aggregated data rates of $N \times 25$ Gbps for each ONU [17]. In addition, NG-EPONs adopt wavelength-fixed transceivers in the ONU instead of wavelength-tunable transceivers. Fig. 2.7 shows an example of the architecture of a 100G NG-EPON.

Increasing the bit rate of a single wavelength above 10 Gbps leads to the reduction

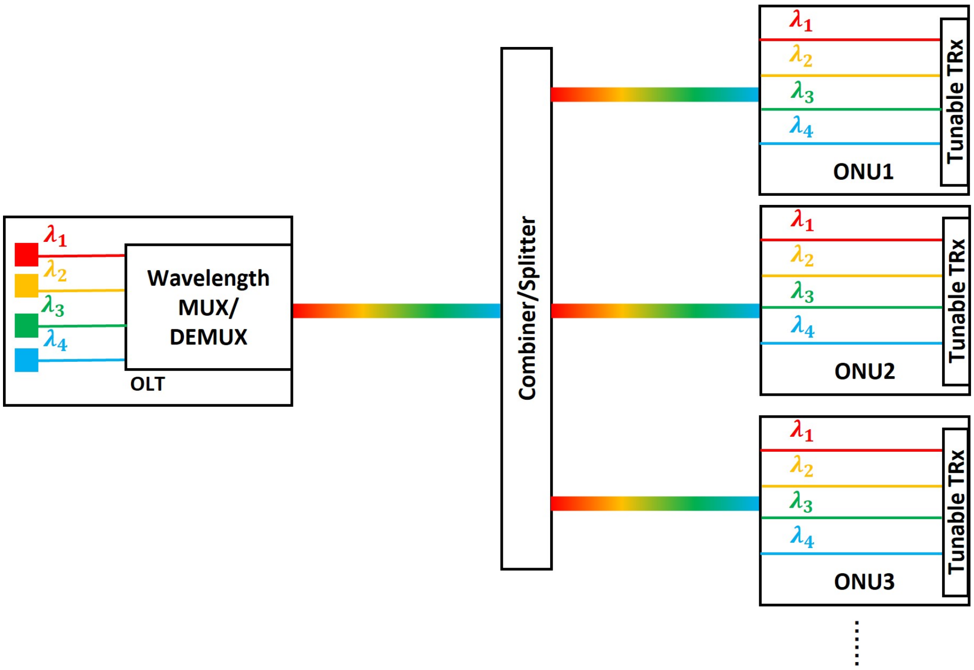

Figure 2.6: The architecture of the NG-PON2

in chromatic dispersion (CD) tolerance and signal to noise ratio (SNR) [25]. The former affects the transmission distance of a fibre while the latter is related to the optical power budget. In addition, a mature ecosystem of low-cost 25G class optics is needed to support the deployment of a large amount of 25 Gbps wavelengths. These are the main challenges in physical layer that NG-EPON needs to overcome [26].

Channel bonding technology is adopted in NG-EPONs so that several wavelengths can be bonded together to form a wavelength channel assigned to a single ONU, achieving higher throughput that is several times of a single wavelength [27, 28, 29, 30]. For instance, two 25 Gbps wavelengths are bonded together to achieve a total peak bit rate of 50 Gbps for an ONU in the 50G NG-EPON. The dynamic wavelength and bandwidth assignment scheme (DWBA) will be adopted in the OLT to efficiently assign bandwidth on several wavelengths to the upstream transmission of an ONU simultaneously. Moreover, the NG-EPON standards are not the standards for a single generation but for multiple generations; specifically, 25G is the first generation, 50G is the second generation and 100G is the third generation [25]. Naturally, subsequent generations after 100G will also be developed. The multiple generations of NG-EPONs must be coexist on the same network and share the ODN elements (i.e., trunk fibres) to achieve low network construction cost.

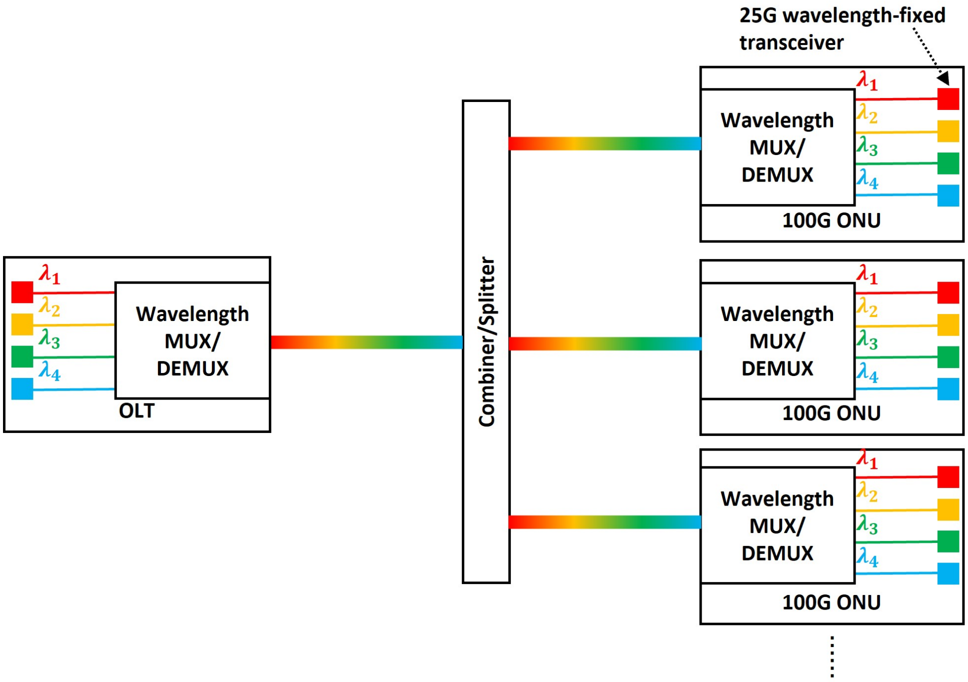

Figure 2.7: The architecture of an 100G NG-EPON

2.2.3 Upstream transmission in NG-EPONs

Efficient dynamic assignment of wavelengths and bandwidth to ONUs for upstream transmission is one challenge of NG-EPONs [20]. The channel bonding technology adopted in NG-EPONs allows several wavelengths being assigned to an ONU at the same time, leading to the DBA/DWBA schemes designed for NG-PON2 cannot be applied in NG-EPONs (since only one wavelength can be assigned to an ONU in NG-PON2). The upstream transmission of an 100G NG-EPON is shown in Fig. 2.8. We can see that the frames in an ONU can be flexibly assigned in 1, 2 or 4 wavelengths to achieve the 25, 50 or 100 Gbps bit rates for an ONU. The DWBA scheme for NG-EPONs dynamically assigns resources to ONUs through GATE and REPORT messages in a round-robin manner, which is same as the conventional DBA/DWBA schemes. When an ONU is heavily loaded, assigning upstream transmissions to too few wavelengths leads to high queuing delay and transmission delay, which seriously affects the network performance. In this case, multiple wavelengths are assigned to the ONU to reduce the upstream delays. Moreover, when an ONU is lightly loaded, assigning the upstream transmissions to multiple wavelengths requires excessive guard time, which wastes upstream bandwidth. In this case, an appropriate number of wavelengths should be assigned for the ONU to avoid guard time waste. Therefore, an efficient DWBA scheme must be developed to reduce bandwidth waste and improve upstream transmission latency in NG-EPONs.

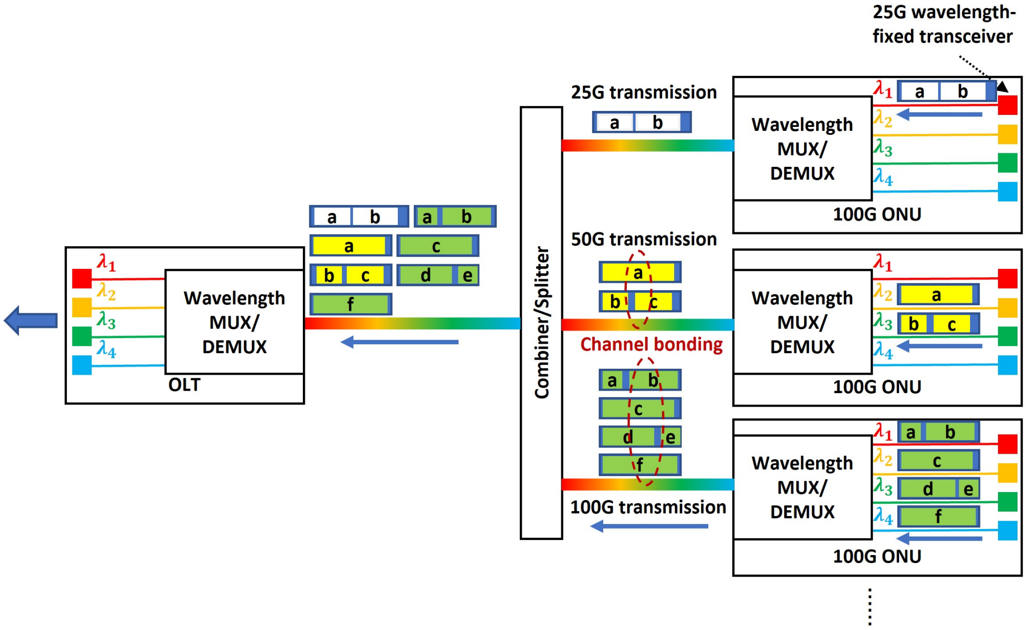

Figure 2.8: The upstream transmission in an 100G NG-EPON

2.3 Mobile x-haul transmission based on PONs

Over the past decades, various researchers have continuously developed cost-effectively PONs to meet future fixed bandwidth requests. However, according to the Cisco forecast highlight report [31], global mobile traffic made up 17% of total IP traffic in 2021, up from 7% from 2016. In contrast, global fixed/wired traffic made up 37% of total IP traffic in 2021, down from 51% in 2016. Moreover, global mobile data traffic is increasing twice as fast as fixed traffic. With the rapid increase in the amount of mobile traffic, potential PON applications that support mobile access services have received considerable attention. For instance, fifth-generation (5G) mobile networks require massive deployment of small cells that can be installed on lamp posts, small towers and other indoor or outdoor locations. In this case, transport networks based on PONs are considered to be more cost-effective and more flexible than networks based on P2P fibres. Specifically, the cost of deploying optical fibres for 5G x-haul transmission using PONs can be reduced by 65% to 95% according to a study by FTTH Council Europe [32]. Furthermore, PONs have been widely deployed in many regions and countries, including Europe, China, Japan, and the United States, to provide access services for fixed traffic. These existing PONs support considerable FTTB/FTTH fibre infrastructure that can be used by mobile network operators (MNOs) to further decrease deployment costs. Therefore, PON applications in mobile x-haul transmission have attracted considerable attentions to address the continuously increasing amount of mobile traffic in the past few years.

2.3.1 Cloud/centralized radio access network

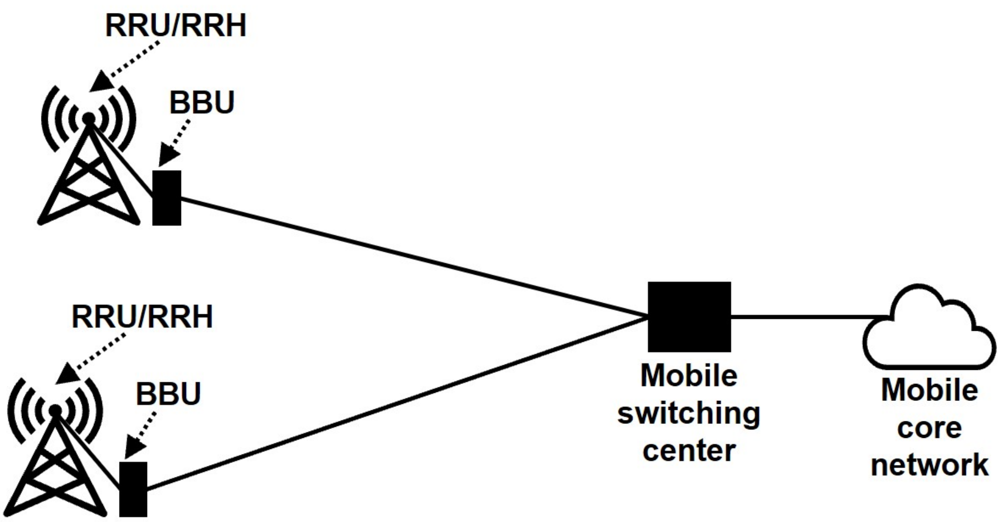

Figure 2.9: D-RAN architecture.

Cloud/centralized radio access network (C-RAN) [33] is the most promising architecture designed for mobile access networks that evolved from distributed radio access network (D-RAN) [34]. In mobile communications, the radio access network (RAN) consists of a lot of base stations (BSs) that provides connection between mobile devices and the core network. The BS contains two components named as the baseband unit (BBU) and the remote radio unit/head (RRU/RRU), respectively. The former performs most of the digital processing and the latter handles the analog processing. In the D-RAN architecture, the BSs are deployed at the cell tower, where the BBU is deployed at the device room near the cell tower and the RRU/RRH is deployed near the antennas which is on the top of the cell tower, as shown in Fig. 2.9. While the BBU and the RRU/RRH are decoupled in the C-RAN architecture, as shown in Fig. 2.10. Specifically, the RRU/RRH is still placed at the top of the cell tower, while the BBUs from different BSs are placed at the same device room near (referred as the BBU pool/BBU hotel) and the computing resources of all BBUs can

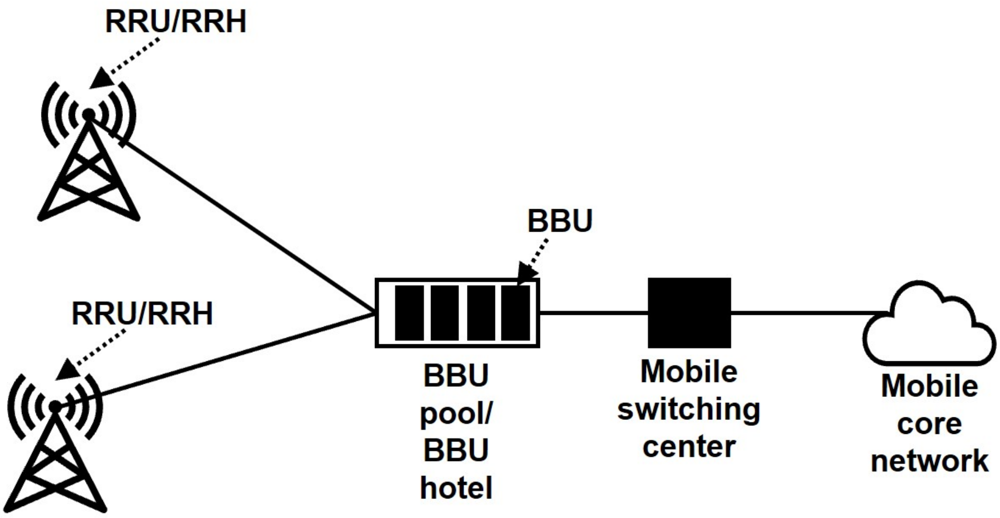

Figure 2.10: C-RAN architecture.

be shared. Due to the aggregation of BBUs, C-RAN architecture enables achieving considerable reduction in power consumption, footprint and management complexity. The BBU and the RRU/RRH in the C-RAN architecture is connected via Common Public Radio Interface (CPRI) [35] and Open Radio equipment Interface (ORI) [35].

2.3.2 Mobile x-haul transmission

As shown in Fig. 2.11, a 5G mobile network based on a C-RAN architecture consists of the new core (NC), the central unit (CU), the distributed unit (DU) and the radio unit (RU) [37]. In the 5G mobile network, the RRU/RRH is renamed as the RU, the BBU is disaggregrated into the DU and the CU and the NC is the core network of 5G mobile network. The transmission between the NC and the CU is known as "mobile backhaul transmission", the transmission between the CU and the DU is known as "mobile midhaul transmission", and the transmission between the DU and the RU is known as "mobile fronthaul transmission (MFH)".

The MFH is based on a higher-layer level split (HLS) interface, which has higher tolerance for latency (a few milliseconds) and requires less bandwidth [22]. The mobile backhaul transmission consists of wired backhaul based on fibre optics and wireless backhaul based on microwave. The latency and bandwidth requirements of wired backhaul are similar to those of the midhaul. Therefore, backhaul and midhaul transmission can be supported by TDM-PONs in a low-cost manner.

The mobile fronthaul (MFH) transmission is based on a lower-layer split (LLS) interface [22] and has a stringent latency requirement ($\leq 100\mu$s [38] or $\leq 250\mu$s

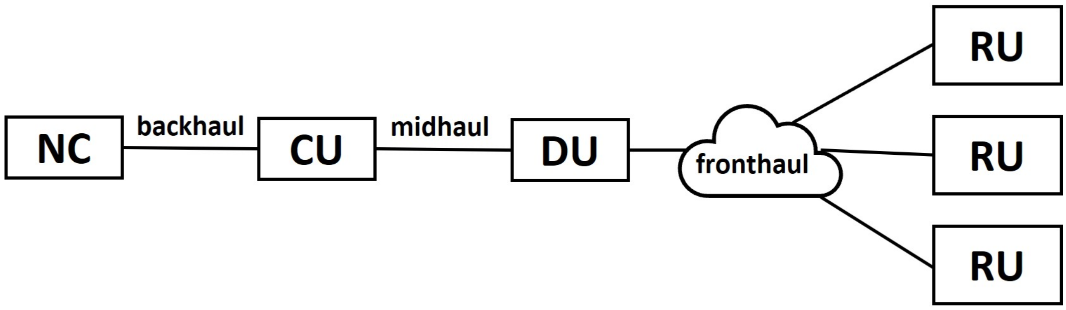

Figure 2.11: A 5G mobile network based on C-RAN architecture.

[39]). However, TDM-PONs have a larger latency (approximately 2 milliseconds) in upstream transmission due to the adoption of DBA schemes. To satisfy the stringent latency requirements of MFH transmission, the ITU-T has begun to standardize the "cooperative DBA" (CO-DBA) scheme, by which the OLT is precisely informed when the RU needs bandwidth [40]. Moreover, MFH transmission requires huge bandwidth, which leads to the need for very high-speed PONs (50 Gbps and up) to achieve an acceptable split level (1:4 split and up) [22]. TWDM-PONs, such as NG-PON2, are expected to satisfy the bandwidth requirement of MFH transmission by stacking several 10 Gbps wavelengths. Increasing the bit rate of a single channel to 50 Gbps or greater is too costly for current components ecosystem [19].

2.3.3 Mobile fronthaul transmission based on time-division duplex

Time-division duplex (TDD) [41] and frequency-division duplex (FDD) [42] are two kinds of duplex technologies adopted in mobile networks. For TDD, both uplink and downlink transmission use the same channel, while transmitters and receivers operate at different time periods, allowing different time periods on the same frequency band to be used for transmitting or receiving data, as shown in Fig. 2.12. For FDD, the uplink and downlink transmission occur in two independent channels, allowing data to be transmitted or received simultaneously and continuously, as shown in Fig. 2.13. TDD technology is more likely to be widely adopted in future mobile networks because TDD approach is more suitable for dense environments with low-power nodes, while FDD approach is not suitable in these environments [43]. Therefore, the MFH requests are assumed to be TDD-based in the future mobile networks.

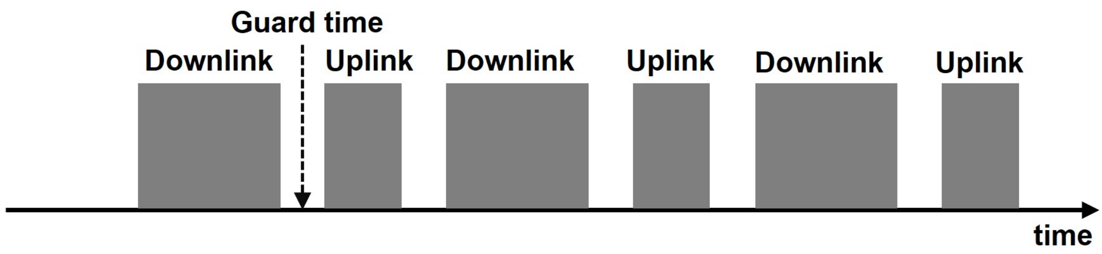

Figure 2.12: Time-division duplex.

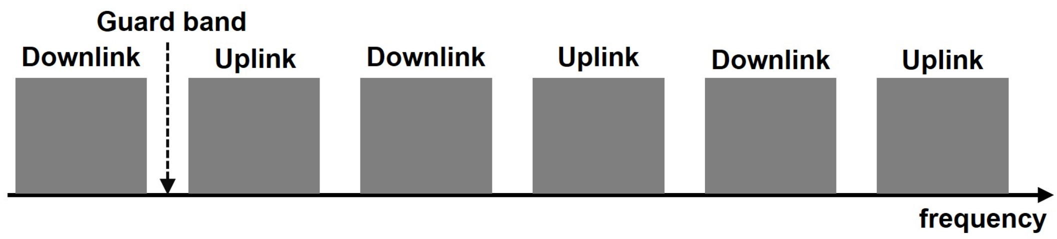

Figure 2.13: frequency-division duplex.

2.3.4 Fixed-mobile convergence access network

Fixed-mobile convergence (FMC), in which fixed and mobile services are offered by a single network, has been emerging in access networks, to further save the total cost of ownership for network providers. J. Kani *et al*. proposed the "access-network slicing" architecture [44], which adopts access elements with higher flexibility to develop a common access platform that is compatible with different services and requests. Moreover, the core network is sliced to provide services for both fixed and mobile transmission through the software-defined networking (SDN)/network functions virtualization (NFV) architecture [45]. X. Liu *et al*. proposed a "futuristic flexible-PON" architecture [46], in which virtual PON (VPON) units as well as software-based routing and capacity sharing are implemented to support various services, including FTTB/FTTH and mobile x-haul. FMC access networks not only provide high-speed, high-capacity and high-flexibility access services for fixed and mobile transmission, but also save the total cost of ownership for network providers.

Fig. 2.14 shows the architecture of an FMC access network, where both mobile and fixed access services can be provided in a single network. In this case, ONUs are connected to RUs or wired devices and the OLT is connected to a swith/aggregator. Moreover, the FMc access network introduce a control plane that is based on SDN technology to flexibly accommodate various services and requirements. The OLT and the switch/aggregator must be NFV infrastructure (NFVI) based on general purpose servers with the necessary factions are modularized [44]. Furthermore, considering the strict latency requirement and the huge bandwidth demands of MFH transmission, the FMC access network is based on the TWDM-PONs.

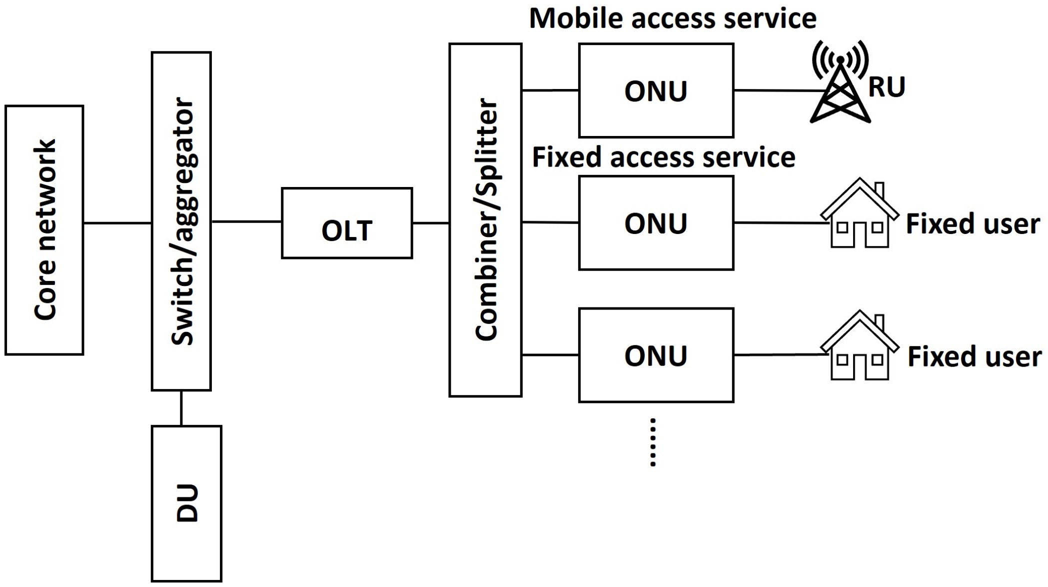

Figure 2.14: The architecture of a FMC access network.

2.3.5 Two types of wavelength multiplexing applied in FMC access network

Two types of wavelength multiplexing are applied in the FMC access network: the mobile transmission is separated from the fixed transmission in the wavelength domain (hereafter referred to as the separate architecture), or the fixed and the mobile transmission share all wavelengths (hereafter referred to as the sharing architecture).

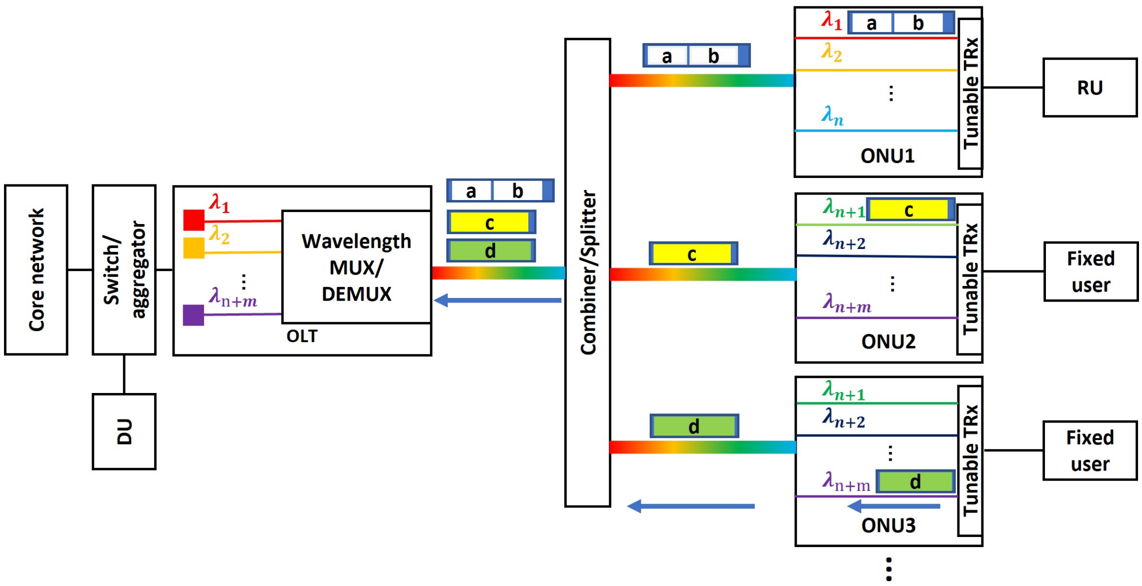

Figure 2.15: "Separate Architecture": fixed and mobile transmission utilize their respective wavelengths.

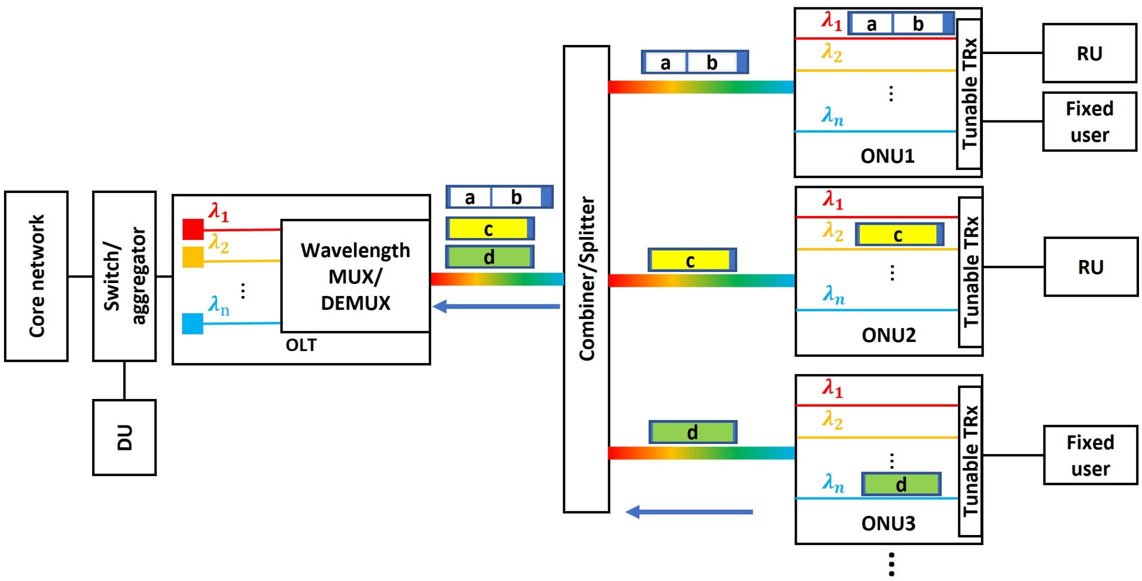

Figure 2.16: "Sharing Architecture": fixed and mobile transmission utilize their respective wavelengths.

The separate architecture based on NG-PON2 is illustrated in Fig. 2.15, where the mobile transmission traffic is able to be assigned to wavelength λ_1 to λ_n while the fixed transmission traffic is able to be assigned to wavelength λ_{n+1} to λ_{n+m}, and the fixed and the mobile transmission are separated in the wavelength domain. In this case, fewer wavelengths are required in each ONU than in the sharing architecture since the fixed and the mobile transmission traffic are assigned to only their respective wavelengths; however, more wavelengths are required in the OLT and more transceivers need to be deployed in the OLT. The sharing architecture based on NG-PON2 is shown in Fig. 2.16, in which the fixed and the mobile transmission traffic are both able to be flexibly assigned to wavelength λ_1 to λ_n. In this case, higher wavelength utilization can be achieved and fewer wavelengths are required in the OLT; however, more wavelengths need to be deployed in each ONU, which increases the cost of the ONUs. Moreover, all transceivers considered in this research are wavelength-tunable since the transceivers deployed in all ONUs must be colorless in NG-PON2.

Chapter 3

Wavelength deployment schemes in terms of ONUs cost and upstream transmission performance in NG-EPONs

As we introduced in Section 2.2.2, the NG-EPON is a promising new standard for EPONs that not only increases the bit rate of a single wavelength to 25 Gbps, but also bonds several 25G wavelengths together to provide an aggregated rate of $N \times 25$ Gbps for a single ONU. In this chapter, we focus on the wavelength deployment scheme selection problem in NG-EPONs. The chapter is organized as follows. In Section 3.1, we introduce the wavelength deployment scheme selection problem. In Section 3.2, we make an overview of previous works and summarize our contribution to this problem. In Section 3.3, we develop an integer linear programming (ILP) model to formulate this problem. In Section 3.4, we propose a heuristic algorithm to efficiently solve the problem in a large-scale network scenario. In Section 3.5, we present numerical analysis for the experiment results. In Section 3.6, we draw conclusions.

3.1 Motivation

Several generations of NG-EPONs must be able to coexist on the same network infrastructure to achieve cost-effective transmission. Therefore, the wavelength deployment scheme selection problem is to determine how many wavelengths should be deployed (prepared) for the coexistence of multi-generation NG-EPONs.

Due to the adoption of channel bonding technology, several 25G wavelengths can be bonded together to achieve high transmission bit rates in NG-EPONs. To make flexible use of all wavelengths, the number of transceivers deployed in the OLT and ONUs is the same as the number of wavelength [47]. On the one hand, deploying fewer 25 wavelengths requires fewer transceivers in an ONU, which decreases the ONU cost, as shown in Fig.3.1(a). However, this approach provides less capacity and less flexibility for upstream transmissions, which leads to worse upstream transmission performance (i.e., higher transmission delay). On the other hand, more 25G wavelengths can be bonded to provide higher peak rates and higher flexibility for upstream transmissions, but this approach leads to the deployment of ONUs with more transceivers, which increases the costs of ONUs, as shown in Fig. 3.1(b). Additionally, OLTs and ONUs of next-generation EPONs can support the transmission of the previous-generation EPONs. Therefore, selecting an appropriate wavelength deployment scheme to balance the upstream transmission performance and ONU cost is very important in the design stage of an NG-EPON.

The architecture of a NG-EPON is designed with backward compatibility [25], which means that we can upgrade NG-EPONs without mass replacement of network infrastructure (i.e., in this research, we consider the ONUs and OLTs). Hence, the ONUs and the OLTs must support the transmission for future EPONs, which means that the structure of ONUs and OLTs must be considered in the network design stage. Moreover, an OLT can be shared by several ONUs, but the number of deployed ONUs is determined by the number of subscribers in an EPON, and ONUs cannot be shared. Even a slight change to the structure of the ONUs substantially affects the construction cost of the whole NG-EPON. Therefore, ONUs are regarded as the most cost-sensitive element in an NG-EPON. To support the flexible assignment between several wavelengths and the coexistence of several generations of NG-EPONs, each ONU has more than one transceiver. By assigning bandwidth in multiple wavelengths, the transmission performance of NG-EPONs can be improved due to the increase in capacity and flexibility. However, the number of transceivers affects the ONU cost. ONUs with excessive transceivers are expensive and greatly increase the construction cost of the overall NG-EPON. Consequently, there is a trade-off between the ONU cost and the upstream transmission performance in an NG-EPON. A suitable wavelength deployment scheme with a cost-effective ONU structure must be developed.

3.2 Previous works and contributions of this research

The previous works mainly focus on how to design an efficient DBA/DWBA scheme to minimize the transmission delay of upstream transmission. In this section, we list several previous works and briefly introduce their research.

In [48], the authors focused on design issues for EPONs and proposed an interleaved polling with adaptive cycle time (IPACT) algorithm for DBA. The IPACT

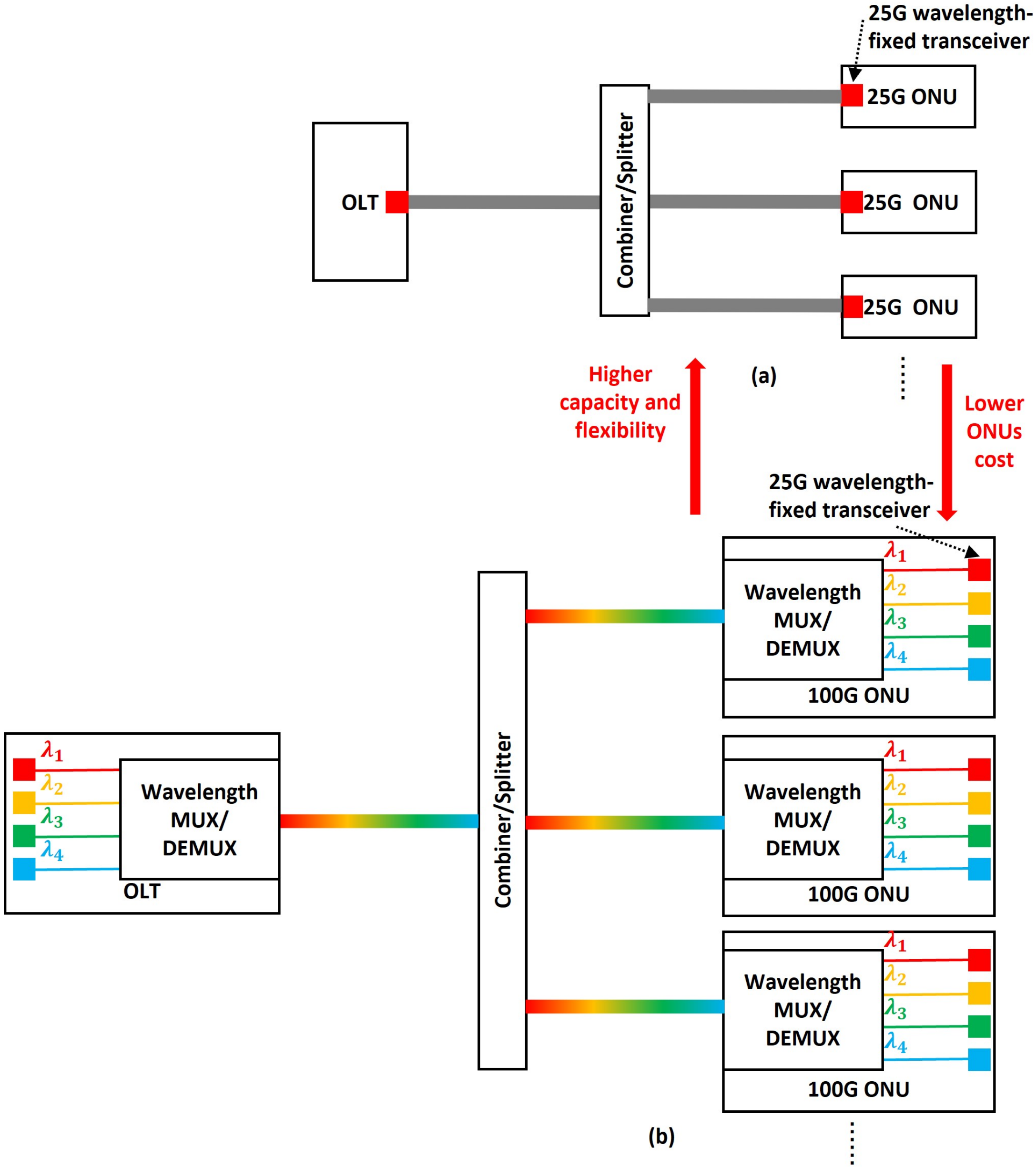

Figure 3.1: Wavelength deployment schemes comparison: (a) deployment of one 25G wavelength; (b) deployment of four 25G wavelengths.

algorithm allows adaptive adjustment of the polling cycle time according to the load of the ONU, leading to fair redistribution of unassigned bandwidth.

In [49], the authors studied three architectures of TWDM-PONs and evaluated their performance. They proposed a water-filling (WF) DWBA algorithm to effectively assign bandwidth and wavelengths for NG-EPONs, and they consider an energy-efficient WF DWBA scheme.

In [50] and [51], the authors aim to improve the upstream performance in NG-EPONs. They propose a first-fit (FF) DWBA algorithm and a flexible wavelength (FW) DWBA algorithm to provide better network performance.

In [52], the authors concentrate on mitigating frame reordering in NG-EPONs. They propose a single channel as possible (SCAP) DWBA algorithm, which is operated through grant readjustment (GP) to reduce frame reordering.

In summary, neither the cost of ONUs nor the wavelength deployment scheme is considered in these works. In this research, we aim to identify the most suitable wavelength deployment scheme for NG-EPONs by analysing the upstream transmission performance considering the cost of ONUs in a static scenario. To the best of our knowledge, our research is the first work to conduct such an analysis.

3.3 Problem formulation

In this section, we analyse the upstream transmission performance considering the ONU cost for different wavelength deployment schemes in a static scenario, and aim to select the most suitable wavelength deployment schemes for NG-EPONs. Different wavelength deployment schemes lead to various numbers of usable wavelengths, which will further influence the network performance and the cost of ONUs. We formulate the trade-off between the ONUs cost and the upstream transmission performance by minimizing the total relative cost, which is the sum of the relative ONU cost and the relative cost transformed by network latency. We develop an ILP model to mathematically formulate the proposed problem. The notation used in this research is defined as follows:

Parameters:

i : index of a request.

N : set of request indices, where $i \in N$.

w : index of a wavelength.

W : set of wavelength indices, where $w \in W$.

a_i : time instant when request i arrives its corresponding ONU.

r_i : required time period (determined by the data size) of request i.

α : relative weight coefficient between the upstream transmission delay and the ONU cost.

M : an integer that is larger than any parameter of variable.

Variables:

W_{max} : maximum index of assigned wavelengths.

$s_{i,w}$: time instant when request i starts transmitting on wavelength w.

$g_{i,w}$: time period granted for request i on wavelength w.

$f_{i,w}$: time instant when request i finishes transmitting on wavelength w.

$k_{i,w}$: a binary value equal to 1 when request i is transmitting on wavelength w and equal to 0 otherwise.

$\xi_{i,j,w}$: a binary value equal to 1 when the start time of request i on wavelength w is earlier than that of request j on wavelength w and equal to 0 otherwise.

d_i : transmission delay of request i.

Objective:

$$\text{Minimize} \quad \alpha \sum_{i=1}^{N} d_i + W_{max}.$$

Constraints:

$$\sum_{1 \le w \le |W|} g_{i,w} \ge r_i \quad \forall i \in N, \tag{3.1}$$

$$f_{i,w} \ge s_{i,w} + g_{i,w} \quad \forall i \in N, w \in W, \tag{3.2}$$

$$g_{i,w} \le M \cdot k_{i,w} \quad \forall i \in N, w \in W, \tag{3.3}$$

$$g_{i,w} \ge k_{i,w} \quad \forall i \in N, w \in W, \tag{3.4}$$

$$W_{max} - M \cdot (k_{i,w} - 1) - w \ge 0 \quad \forall i \in N, w \in W, \tag{3.5}$$

$$s_{i,w} \ge k_{i,w} \cdot a_{i,w} \quad \forall i \in N, w \in W, \tag{3.6}$$

$$s_{i,w} \le M \cdot k_{i,w} \quad \forall i \in N, w \in W, \tag{3.7}$$

$$\xi_{i,j,w} + \xi_{j,i,w} = 1 \quad \forall i \in N, j \in N, w \in W, \tag{3.8}$$

$$f_{i,w} - s_{j,w} \le M \cdot (3 - k_{i,w} - k_{j,w} - \xi_{i,j,w}) \quad \forall i \in N, j \in N, w \in W, \tag{3.9}$$

$$d_i \ge f_{i,w} - k_{i,w} \cdot a_i - M \cdot (1 - k_{i,w}) \quad \forall i \in N, w \in W. \tag{3.10}$$

Constraint (3.1) ensures that the granted time periods for request i in all wavelengths are greater than or equal to the required data size. Constraint (4.2) implies that the completion time of request i should be later than the sum of its start time and its granted time periods in every wavelength. Constraint (4.3) and Constraint (4.4) mean that if request i is assigned to a wavelength (i.e., $k_{i,w} = 1$), then a corresponding time period for this request must be granted in the corresponding wavelength; otherwise, there will be no time period granted for this request. Constraint (4.5) denotes the maximum index of the assigned wavelengths. In Constraint (4.6) and Constraint (4.7), the transmission start time of request i in a wavelength must be later than its arrival time if the request is assigned to this wavelength (i.e., $k_{i,w} = 1$); otherwise, the transmission start time of this request must be 0. Constraint (4.8) and Constraint (4.9) are non-overlapping constraints. According to Constraint (4.8), if two different requests are assigned to the same wavelength, then there must be a priority relationship between their transmission start times, which means only one among $\xi_{i,j,w}$ and $\xi_{j,i,w}$ is equal to 1. Furthermore, in Constraint (4.9), if request i and request j are not assigned to the same wavelength w (i.e., $k_{i,w} = 0$ or $k_{j,w} = 0$), then there is no need to consider the overlap between them. The right side of Constraint (4.9) is a very large positive integer that is always greater than the left part of Constraint (4.9). If request i and request j are assigned to the same wavelength w (i.e., $k_{i,w} = 1$ and $k_{j,w} = 1$), then the non-overlapping constraint must be satisfied. When the start time of request i on wavelength w is earlier than that of request j (i.e., $\xi_{i,j,w} = 1$), then the right part of Constraint (4.9) is equal to 0. Moreover, Constraint (4.9) becomes

$$f_{i,w} - s_{j,w} \leq 0 \quad \forall i \in N, j \in N, w \in W,$$

which ensures that the transmission completion time of request i is earlier than the transmission start time of request j. Therefore, there is no overlap between request i and request j. Otherwise, if the start time of request i on wavelength w is earlier than that of request j (i.e., $\xi_{i,j,w} = 0$), then Constraint (4.9) is always satisfied. Constraint (4.10) denotes that the delays of all requests must be greater than the difference between the latest transmission completion time and the next arrival time.

To achieve the most flexible wavelength assignment in an NG-EPON, wavelength-tuneable transceivers can be deployed in ONUs; however, the cost of wavelength-tuneable transceivers is much higher than the cost of wavelength-fixed transceivers. Clearly, there is a trade-off between network performance and cost. The cost of ONUs is the most sensitive factor in the total NG-EPON construction cost, and deploying wavelength-tuneable transceivers will greatly increase the cost of ONUs. This is contrary to the original intention of NG-EPONs to make economic sense. Hence, all transceivers in the ONUs are wavelength fixed in NG-EPON. The concept of the maximum index of assigned wavelengths rather than the maximum number of assigned wavelengths is used in our ILP formulations, which can substantially reduce the symmetry of the ILP model and improve the computational efficiency.

In this case, we can view the simplest case of the wavelength and bandwidth allocation problem (i.e., only one wavelength can be assigned to requests) as classical scheduling theory [53], which has been proved to be NP-hard [54]. Specifically, the ONUs in an NG-EPON can be regarded as jobs, and the granted time periods for these ONUs can be regarded as the processing times of the jobs. Clearly, the original problem is also NP-hard since it is far more complex than the simplest case. Additionally, the complexity of this ILP model is $O(WN^2)$, where W and N are the number of candidate wavelengths and the number of transmission requests, respectively. Consequently, the computation time of this ILP model increases substantially with increasing numbers of candidate wavelengths and transmission requests.

3.4 Wavelength deployment scheme comparison and selection algorithm

In this section, we propose a WDSCS algorithm to efficiently solve the problem of how to balance the performance of an EPON and ONU cost. Specifically, WDSCS compares the relative costs of all wavelength schemes and selects the most suitable scheme. The central idea of WDSCS is to minimize the total upstream transmission delay for every wavelength scheme. Note that the total upstream transmission delay consists of two components: transmission delay and queuing delay [48]. Transmission delay is determined by the data size of the upstream transmission requests and the upstream transmission rate, while queuing delay depends on the waiting time between the request arrival time and the transmission start time. Therefore, two methods are applied in WDSCS. One is to increase the upstream transmission rate so that every upstream transmission request is assigned equally among all wavelengths in every wavelength deployment scheme in WDSCS. The other approach is to properly schedule the upstream transmission requests to reduce the queuing delay. The total upstream transmission delay is affected by the scheduling scheme since the transmission delays for every wavelength deployment scheme are determined if the upstream requests are given a priori.

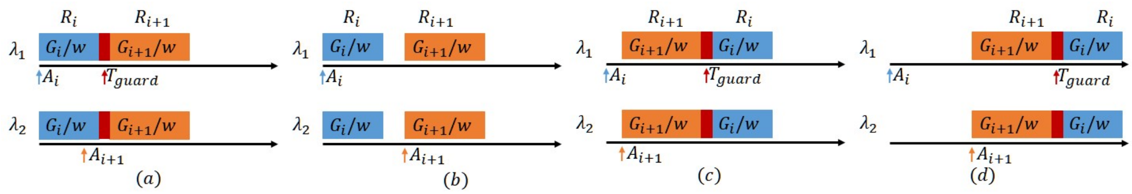

Figure 3.2: Scheduling of two adjacent requests R_i and R_{i+1}: (a) R_i is assigned earlier when R_{i+1} arrives before R_i finishes transmitting; (b) R_i is assigned earlier when R_{i+1} arrives after R_i finishes transmitting; (c) R_{i+1} is assigned earlier when R_{i+1} arrives before R_i can finish transmitting; (d) R_{i+1} is assigned earlier when R_{i+1} arrives after R_i can finish transmitting.

The pseudocode of WDSCS is presented in Algorithm 1. The upstream transmission request volume, including the arrival time and the traffic size, is given. First, we sort all requests in increasing order of arrival time A_i of each request R_i and calculate the granted time period G_i in a wavelength (by equally assigning every upstream transmission request among all wavelengths). $Rate_\lambda$ presented in Line 3 is the upstream transmission rate of a wavelength. In the next step, we start scheduling for different wavelength schemes. If the arrival time A_i of a request R_i is earlier than the transmission completion time T_c of a previously assigned request in Line 7 (T_g is the guard time), then R_i is added to a waiting pool denoted by WP, and the request with the minimum data size R_{min} in the waiting pool is selected in Line 9. If A_i is later than T_c, two cases arise. When WP is empty, we assign R_i if G_i minus G_{i+1} divided by the wavelength number w is less than the value of A_{i+1} minus A_i multiplied by 2, as shown in Line 12; otherwise, R_i is added to WP. We follow this approach because if we assign R_i earlier than R_{i+1}, then the delay of R_i is G_i/w, and the delay of R_{i+1} is $(G_i + G_{i+1})/w - (A_{i+1} - A_i)$ if R_{i+1} arrives before R_i finishes transmitting or G_{i+1}/w if R_{i+1} arrives after R_i finishes transmitting. The total delay of R_i and R_{i+1} is $(2G_i + G_{i+1})/w - (A_{i+1} - A_i) + T_g$ or $(G_i + G_{i+1})/w$, as shown in Fig. 3.2(a) and Fig. 3.2(b), respectively. Otherwise, if we assign R_i later than R_{i+1}, then the total delay of R_i and R_{i+1} is $(G_i + 2G_{i+1})/w + (A_{i+1} - A_i) + T_g$, as shown in Fig. 3.2(c) (R_{i+1} arrives before R_i can finish transmitting) and Fig. 3.2(d) (R_{i+1} arrives after R_i can finish transmitting). Clearly, $(G_i + 2G_{i+1})/w + (A_{i+1} - A_i) + T_g$ is always greater than $(G_i + G_{i+1})/w$, so we assign R_i earlier than R_{i+1} in this case. Another case is to compare $(2G_i+G_{i+1})/w-(A_{i+1}-A_i)+T_g$ and $(G_i+2G_{i+1})/w+(A_{i+1}-A_i)+T_g$. If the former is less than the latter, then $(G_i - G_{i+1})/w \leq 2 * (A_{i+1} - A_i)$ can be observed, and we assign R_i earlier than R_{i+1}, as shown in Line 15; otherwise, we add R_i to WP. If the interval between A_i and T_c is greater than that of the minimum request G_{min} and WP is not empty, as shown in Line 19, then R_{min} can be "inserted" between the current assigned request and R_i without affecting the queuing delay of R_i. In this case, R_{min} is equally assigned among every wavelength before R_i and removed from WP; then, a new minimum request is selected from the remaining requests in WP, as shown in Lines 20 to 23. R_i continues to be compared with the minimum request until the interval between A_i and T_c is less than G_{min}. Then, we compare the data size of R_i and R_{min}. In Lines 24 to 31, we schedule R_i and R_{min} according to the previously described method. After traversing all requests, the requests remaining in WP are assigned in increasing order of their data size in Lines 33 to 36. Finally, we calculate the total delay and the relative cost for every wavelength scheme.

Algorithm 1 WDSCS algorithm

Input: R,A

Sort all requests in increasing order of the arrival time A_i of each request R_i
for R_i in R **do**
    $G_i = R_i \;/\; Rate_\lambda$
end for
for w in W **do**
    Initialize $WP = \varnothing, T_c = 0$
    for R_i in R **do**
        if $A_i < T_c + T_g$ **then**
            Add R_i to WP, and find the minimum
            R_{min} of WP;
        else
            if $WP = \varnothing$ **then**
                if $(G_i - G_{i+1})/w \leq 2 * (A_{i+1} - A_i)$ **then**
                    Assign the ith request R_i, calculate the
                    delay D_i, update T_c;
                else
                    Add R_i to WP, and find the
                    minimum R_{min} of WP;
                end if
            else
                while $A_i - T_c - T_g \geq G_{min}$ & $WP \neq \varnothing$ **do**
                    Assign the request R_{min}, calculate the
                    delay D_{min}, update T_c, remove R_{min}
                    from WP, find the new minimum
                    R_{min} of WP;
                end while
                if $(G_{min} - G_i)/w \leq 2 * (A_i - T_c)$ **then**
                    Assign the request R_{min}, calculate the
                    delay D_{min}, update T_c, remove R_{min}
                    from WP, add R_i to WP, find the
                    new minimum R_{min} of WP;
                else
                    Assign the ith request R_i, calculate the
                    delay D_i, update T_c;
                end if
            end if
        end if
    end for

39: **while** $WP \neq \varnothing$ **do**
40: Assign the request R_{min}, calculate the
41: delay D_{min}, update T_c, remove R_{min}
42: from WP, find the new minimum
43: R_{min} of WP;
44: **end while**
45: $D_w \leftarrow \sum_{i=1}^{N} D_i$
46: $C_w \leftarrow \alpha^2 \cdot D_w + w$
47: **end for**
Output: D_w, C_w

The time complexity is mainly determined by the iterations of Lines 4 to 36. Since the maximum number of iterations of Line 19 is $O(log(N-2))$, there are N-2 requests in the waiting pool, and the interval between the arrival time of the last request and the completion time of the assigned request is sufficient to hold all N-2 requests. Therefore, the total time complexity can be expressed as $O(WNlog(N-2))$, where W denotes the number of wavelength schemes, and N denotes the total number of requests.

3.5 Numerical analysis

In this section, we present the results of our experiments. All experiments are conducted in a static scenario in this research, which means that the data size and the arrival time of all upstream traffic volumes are given a priori. Therefore, there is no need to consider the round-trip times of the GATE and REPORT messages. An EPON based on the tree topology is adopted in our experiments; moreover, the OLT and the ONUs support a bitrate of 25 Gbps per wavelength. We consider 8 ONUs connected to an OLT for simplicity in a small network scenario. Moreover, 128 ONUs connected to an OLT are considered to verify the NG-EPON performance for future heavy traffic volumes in a large network scenario. The parameters used in this research is shown in Table. 3.1

Table 3.1: Parameters for experiments

Parameter	Value
Upstream link bandwidth	25 Gbps per wavelength
Number of ONUs	8 / 128
Number of requests	50 / 800
Ethernet frames	64 - 1518 bytes
weight coefficients	α

Self-similar traffic is generated in our simulation to reflect the network traffic in reality. According to [55], self-similar traffic can be obtained by aggregating multiple sub-streams with alternating Pareto-distributed ON/OFF periods.The Hurst parameter is 0.8 [56], and $\mathrm{E}[t_{ON}]/(\mathrm{E}[t_{ON}] + \mathrm{E}[t_{OFF}]) = 1/10$ is the load of each sub-stream, where $\mathrm{E}[t_{ON}]$ and $\mathrm{E}[t_{OFF}]$ denote the expected values of the ON and OFF periods, respectively [57]. Moreover, Ethernet frames are uniformly generated from 64 to 1518 bytes, and a guard time of 1s is assumed to distinguish upstream requests transmitted by different ONUs. The inter-frame gap between adjacent Ethernet frames in an upstream request is 12 bytes, and the preamble of the frame is 8 bytes. The offered network load is denoted as the ratio between the total traffic requests generated by all ONUs per second and the maximum capacity of all wavelengths. Different relative weight coefficients α lead to various relative costs of NG-EPONs; thus, various wavelength deployment scheme are suggested. In our simulation experiments, α ranges from 0.02 to 0.06 in the small network scenario, and from 0.005 to 0.015 for the large case. For every offered network load, we generate 50 upstream requests as the input for the small scenario and 800 requests for future heavy traffic volumes. To improve the accuracy of the simulation, we perform 50 iterations with independent inputs and take the average as the final result. We use the GUROBI Optimizer v9.1.1 [58] to solve the ILP models proposed in Section III. Since the problem is NP-hard, the ILP model is difficult to solve completely within a reasonable time in some cases. Therefore, the execution time of the proposed ILP models is bounded at 3600 s. We use a computer with an Intel 6-core 12-thread 3.7 GHz CPU and 32 GB of memory to conduct all simulation experiments in a Microsoft Windows 10 environment. All simulation programs are developed using Python v3.8.

In Fig. 3.3, we show the average of the maximum numbers of required wavelengths under different network loads and relative weight coefficients α simulated by ILP and our proposed WDSCS. The results of both WDSCS and ILP indicate that the maximum numbers of required wavelengths increase with an increase in α from 0.02 to 0.06 for a given network load. This is because a greater value of α means that network operators pay more attention to latency, and more wavelengths are required to be deployed for increasing the capacity to reduce the transmission delay. In addition, there is a tendency for the maximum numbers of required wavelengths to increase with increasing the offered network load. Since a higher offered load results in a higher queuing delay when the number of upstream requests is fixed, to reduce the total relative cost, upstream requests are assigned to more wavelengths to decrease the transmission delay, which leads to a increase of the maximum number of required wavelengths. From the simulation results, WDSCS performs similarly to ILP, and the gap between WDSCS and ILP is up to 3.6% when the offered network load is 0.5 and α is 0.04. Additionally, the average gaps among all offered network loads are 1.1, 1.4, 1.4, 1.4 and 0.9% when α increases from 0.02 to 0.06; therefore our proposed WDSCS can efficiently select a suitable wavelength deployment scheme.

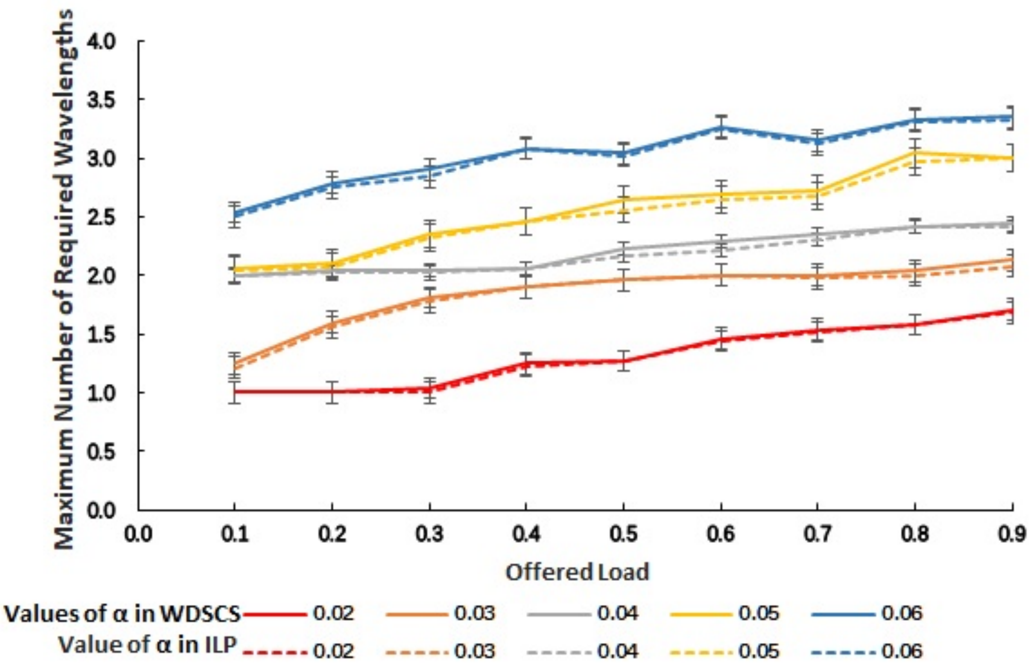

Figure 3.3: Average of the maximum number of required wavelengths for ILP and WDSCS.

Fig. 3.4 shows the average upstream transmission delays of ILP and WDSCS (WDSCS uses the same number of wavelengths as ILP to calculate the upstream transmission delay). For a given α, the average upstream transmission delays of both ILP and WDSCS vary only slightly with the change in the offered network load, especially when the offered network load is between 0.3 and 0.7. When α is 0.03, the delays even decrease as the network load increases from 0.1 to 0.3. This is mainly due to the joint influence of the increasing offered network load and the deployment of multiple wavelengths. An increase in the offered network load leads to a high transmission delay, which results in a tendency to increase the total relative cost; thus, multiple wavelengths are recommended to reduce the transmission delay corresponding to this tendency. On the other hand, the average upstream transmission delays of both methods decrease substantially with an increase in α for a given network load. As mentioned in the previous paragraph, the number of deployed wavelengths increases substantially with increasing α. Lower upstream transmission delays can be provided given the deployment of more wavelengths. Moreover, WDSCS has a delay that is at most 4.2% greater than that of ILP when the offered network load is 0.7 for an α of 0.04. Therefore, WDSCS can efficiently schedule and assign upstream requests to minimize the transmission delay.

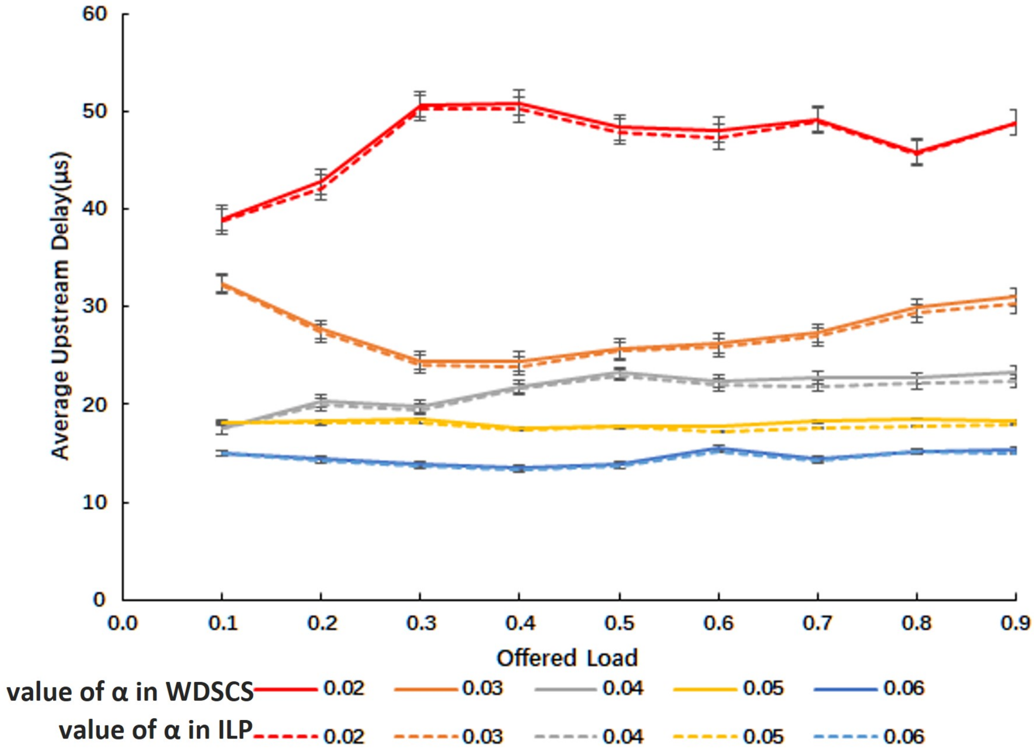

Figure 3.4: Average upstream transmission delays of ILP and WDSCS.

In Fig. 3.5, we compare the average upstream transmission delays calculated by our proposed WDSCS and several algorithms proposed in previous works when the number of deployed wavelengths is 4. Fig. 3.5 shows that WDSCS performs better than modified-IPACT, WF-DBA and FF-DBA under any network loads. FF-DBA gives the highest upstream transmission delay because the transmission rate of each upstream traffic is the lowest in this case. Modified-IPACT and WF-DBA outperform FF-DBA since they assign upstream traffic in all wavelengths to provide a maximum transmission rate; however, the queuing delay is not considered in these approaches. Our proposed WDSCS provides a maximum transmission rate and a low queuing delay, so WDSCS achieves the lowest upstream transmission delay. In addition, the difference increases since more queuing occurs under high network load.

Jitter analysis (the standard deviation of delay) is shown in Fig. 3.6 to further evaluate modified-IPACT, WF-DBA, FF-DBA and WDSCS. All approaches provide low jitter under low network load since all upstream traffic can be transmitted with low delays in this case. The jitter of Modified-IPACT, WF-DBA and FF-DBA significantly increases when the network load is over 0.5, because several upstream traffic are highly delayed due to the queuing. When the network load is heavy (over 0.8), the

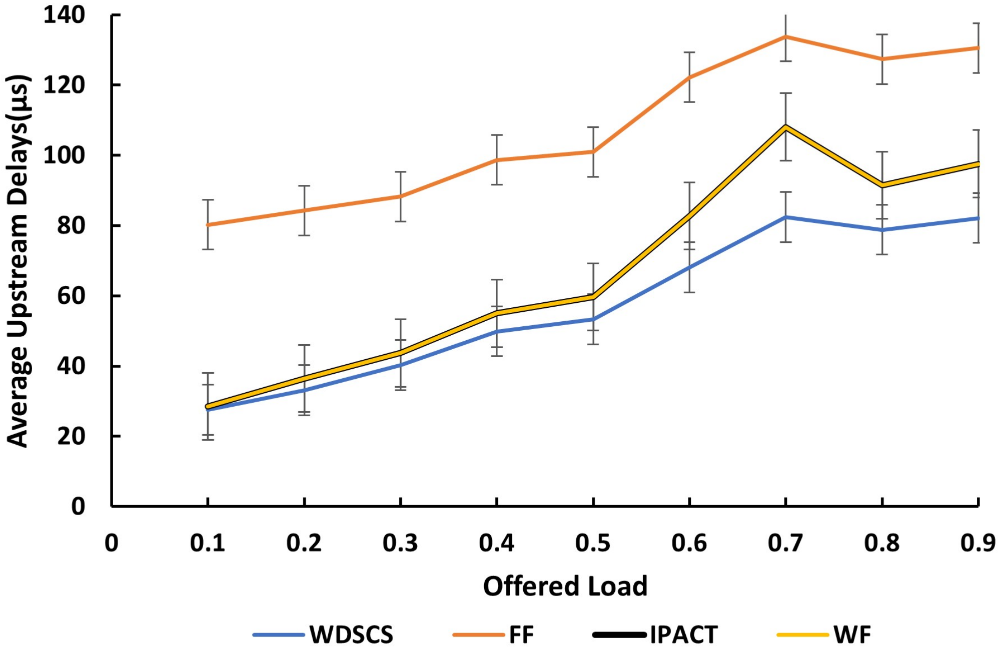

Figure 3.5: Average upstream transmission delays of modified-IPACT, WF-DBA, FF-DBA and WDSCS.

delay of almost all upstream traffic is high because of the queuing; thus the jitter of Modified-IPACT, WF-DBA and FF-DBA decreases when the network load is heavy, which leads to the peaks in the lines of Modified-IPACT, WF-DBA and FF-DBA. Moreover, our proposed WDSCS gives the lowest jitter under all network loads and provides a flat line since it can achieve the lowest upstream transmission delay with low queuing delay.

In Fig. 3.7, we compare the average upstream delays (solid lines) and the required wavelengths (dotted lines) calculated by WDSCS in four cases: the number of deployed wavelengths is (a)one, (b)two, (c)four, and (d) eight. In Fig. 3.7(a), (b) and (c), α is 0.02, 0.04 and 0.06, respectively. When α is given, as shown in Fig. 3.7(a), the upstream delay first sharply decreases as the number of deployed wavelengths increases for a given network load (more than 0.1); then, the upstream delay varies slightly with increasing number of deployed wavelengths. When the deployed wavelength increases from 1 to 4, the upstream delays are reduced by approximately 1 to 4 times according to the network load; however, when the deployed wavelength increases from 4 to 8, the upstream delays are almost equal (solid lines coincide). This can be explained by analysing the required wavelengths. When deploying only a few wavelengths, the upstream delay is very high since the capacity is not sufficient.

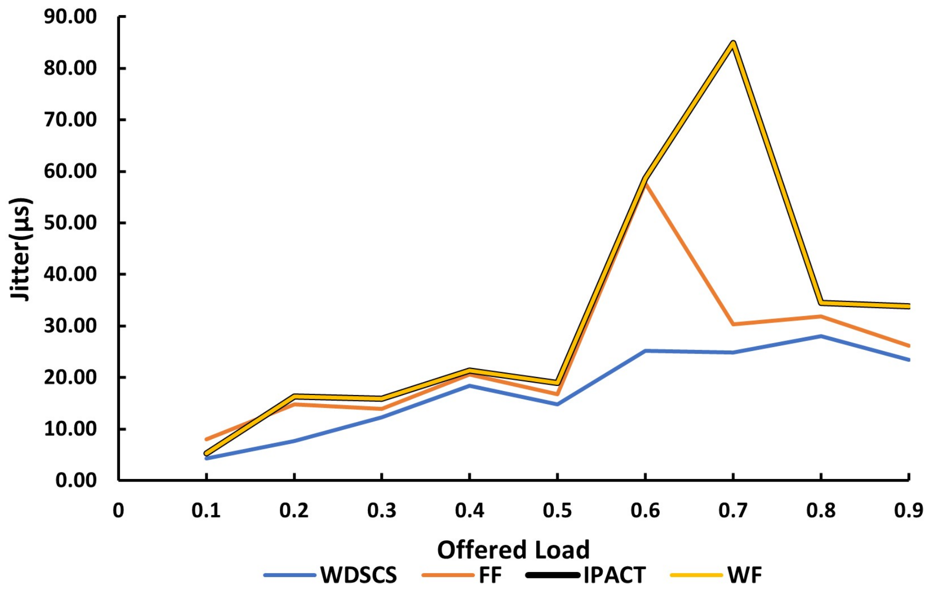

Figure 3.6: Jitter of modified-IPACT, WF-DBA, FF-DBA and WDSCS.

In this case, the gain of reducing the upstream delay is higher than low ONU cost; thus, more wavelengths are required to increase the capacity, and the upstream transmission delays significantly decrease with the deployment of more wavelengths. On the other hand, when multiple wavelengths are deployed, the capacity is sufficient. In this case, cost-effective ONUs are preferred to low upstream delays; therefore, no more wavelengths are required, and the upstream transmission delays change slightly. Similar results can be yielded from Fig. 3.7(b) and Fig. 3.7(c). Moreover, different α values lead to different required wavelengths, and various most suitable wavelength deployment schemes can be selected. Therefore, our proposed WDSCS can balance the ONU cost and the upstream transmission performance of NG-EPONs to choose a cost-effective wavelength deployment scheme.

Fig. 3.8 shows the average upstream delays (solid lines) and the required wavelengths (dotted lines) calculated by WDSCS of several wavelength deployment schemes (the numbers of deployed wavelengths are 2,4,8 and 16, respectively) for future heavy traffic volumes. The offered network load is up to 0.7 since the network load is usually less than 0.7 in real networks. The results is similar to Fig. 3.7. Therefore, network operators can choose the most suitable wavelength deployment scheme to ensure the low cost of ONUs and good performance by our proposed WDSCS when designing the NG-EPONs in the future. In addition, network operators can provide QoS-based service by setting different relative weight coefficients α in future network scenarios.

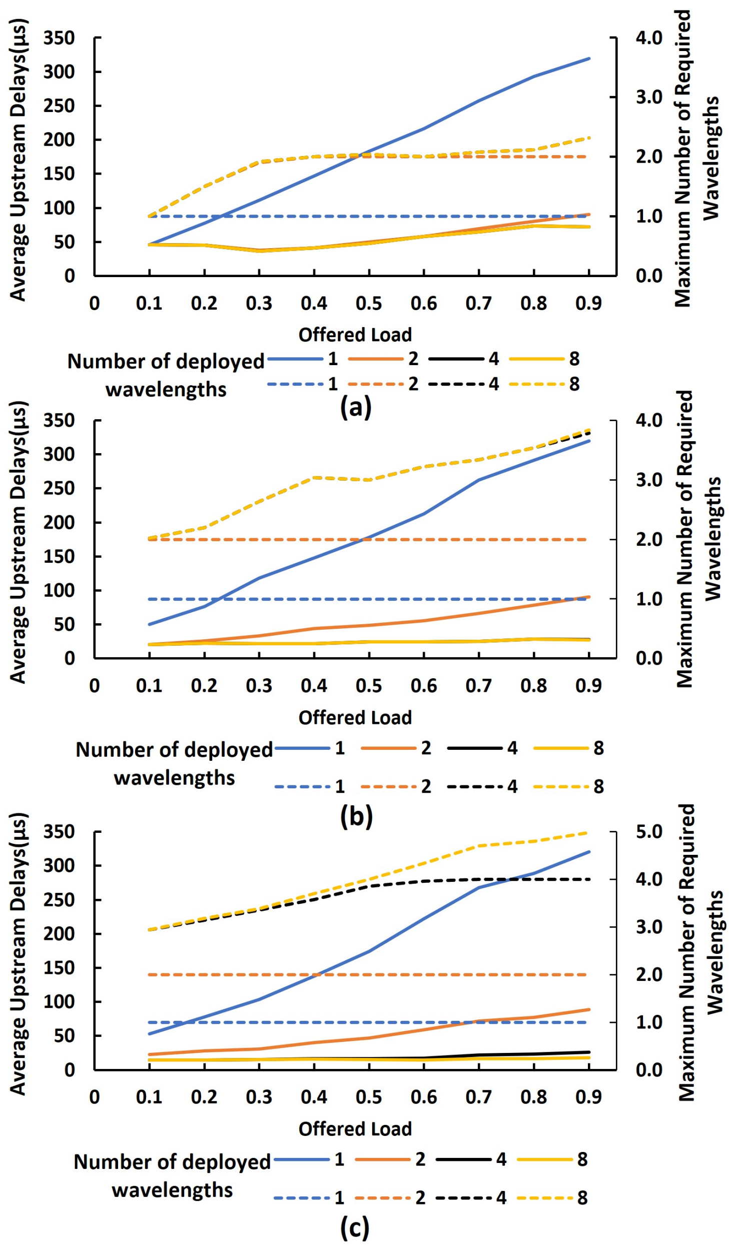

Figure 3.7: Average upstream delays and required wavelengths of different wavelength deployment schemes: (a) $\alpha = 0.02$; (b) $\alpha = 0.04$; (c) $\alpha = 0.06$.

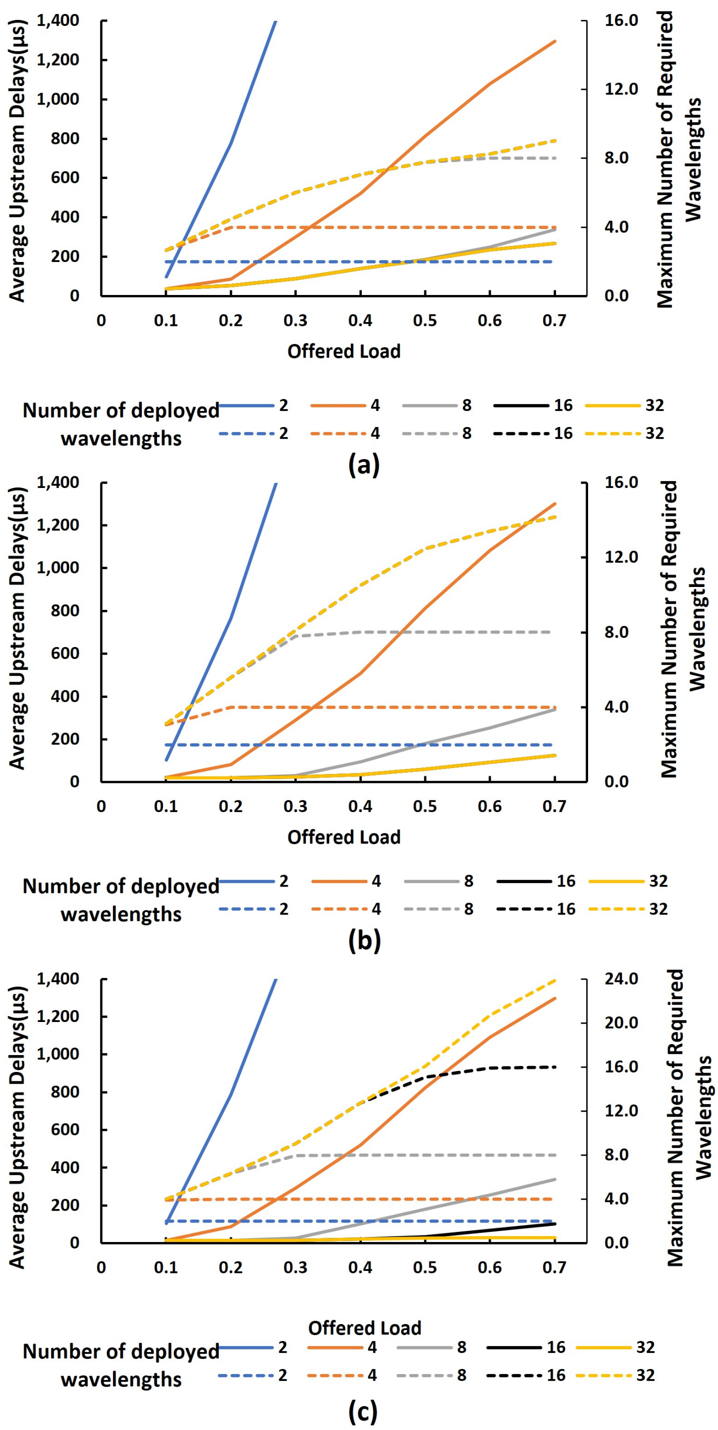

Figure 3.8: Average upstream delays and required wavelengths of different wavelength deployment schemes for future heavy traffic volumes: (a) $\alpha = 0.005$; (b) $\alpha = 0.01$; (c) $\alpha = 0.015$.

3.6 Conclusion

In this research, we discuss the design of ONUs and wavelength deployment schemes in an NG-EPON. Since ONUs are the most cost-sensitive element, we develop an ILP model to minimize the total relative cost in a static EPON scenario. To solve the problem efficiently, we propose a heuristic WDSCS algorithm. From our simulation results, we can see that our proposed WDSCS performs similarly to the ILP model in terms of the number of deployed wavelengths and the average upstream transmission delay. A comparison of between WDSCS and WF shows that our proposed algorithm performs better in minimizing the upstream transmission delay. The upstream transmission delays and required wavelengths of different wavelength deployment schemes are analysed in small network scenarios, and WDSCS can effectively choose the most suitable wavelength deployment schemes to find a satisfactory balance between network performance and ONU cost. Moreover, simulations of future heavy traffic volume scenarios are provided as a reference for future NG-EPONs design. In this way, network operators can select the most cost-effective NG-EPON design considering both the ONU cost and the upstream transmission performance. QoS-based service can also be provided to satisfy the different needs of users.

Chapter 4

Wavelength deployment scheme in a fixed-mobile convergence access network

As we discussed in Section 2.3, FMC is emerging in access network to provide access services for both fixed and mobile transmission in a single network. In this chapter, we mainly focus on the problem of how to select the appropriate wavelength deployment scheme in a TWDM-PON-based FMC access network. The chapter is organized as follows. In Section 4.1, we introduce motivation of this research. In Section 4.2, we overview several previous works and summarize our contribution. In Section 4.3, we develop an ILP model to formulate this problem. In Section 4.4, we propose a heuristic algorithm to efficiently solve the problem in a large-scale network scenario. In Section 4.5, we present numerical analysis for the experiment results. In Section 4.6, we draw conclusions.

4.1 Motivation

The mobile x-haul transmission, especially the MFH transmission, has more stringent latency requirement than fixed upstream transmission. The FMC access network must be able to satisfy the latency requirement of both MFH transmission and fixed transmission, which requires the FMC access network to deploy sufficient amount of wavelengths. However, the power consumption and the deployment cost of the total network is related to the number of wavelengths deployed in the network. In this research, we aim at minimizing the number of wavelengths deployed in the FMC access network while satisfying the latency requirement of both MFH and fixed transmission. Moreover, two potential architectures of the FMC access network are introduced, which differ mainly in the wavelength multiplexing solutions for fixed and MFH transmission. We compare the separate architecture and the sharing architecture under different network parameters.

4.2 Previous works and contributions of this research

Many papers in the literature have focused on DBA/DWBA schemes for MFH transmission in TDM-/TWDM-PONs since the conventional online and offline DBA schemes are message-based and cannot satisfy the stringent low-latency requirements.

In [59], the authors proposed a mobile DBA scheme in which the MFH latency requirement is satisfied by utilizing the scheduling information in the wireless domain.

In [60], the authors proposed an MFH DBA scheme based on simple statistical traffic analysis, and an upstream latency under 50 μs was achieved according to the experimental results.

In [61] and [62], the authors proposed efficient DWBA schemes for MFH transmission in TWDM-PON, and they further considered the synchronization timing errors in the TDD-based fronthaul and the data size of a TDD burst is determined by the 5G parameters.

In summary, these authors only considered how to meet the latency requirements of mobile forward transmission in PON, without considering that the FMC access network can satisfy the latency requirement of both fixed and MFH transmission. In this research, we mainly focus on the design problem of a FMC access network based on NG-PON2. We aim at find the minimal number of wavelengths that can support the latency requirement of both fixed and MFH transmission. We also present a comparison between the two potential architectures under different network parameters. To the best of our knowledge, this research is the first work to conduct such research on the FMC access network.

4.3 Problem formulation

In this section, we consider the problem of how to minimize the number of required wavelengths in an FMC access network. To accomplish this, we develop an ILP model. The parameters and variables used in this research are shown below.

Parameters:

i : index of a request.

N : set of request indices, where $i \in N$.

a_i : time instant when request i arrives at its corresponding ONU.

t_i : latency threshold of request i.

r_i : required time period (determined by the data size) for request i.

w : index of a wavelength.

W : set of wavelength indices, where $w \in W$.

M : an integer that is larger than any parameter or variable.

Variables:

W_{max} : maximum index of the assigned wavelengths.

$s_{i,w}$: time instant when request i starts transmitting on wavelength w.

$g_{i,w}$: time periods granted to request i on wavelength w.

$f_{i,w}$: time instant when request i completes transmitting on wavelength w.

$k_{i,w}$: a binary variable that is equal to 1 if request i is transmitted on wavelength w and equal to 0 otherwise.

$\xi_{i,j,w}$: a binary variable that is equal to 1 if request i starts transmitting earlier than request j on wavelength w and equal to 0 otherwise.

d_i : transmission delay of request i.

The wavelength assignment problem is formulated as follows:

Objective:

$$\text{Minimize} \quad W_{max} \tag{4.1}$$

subject to

$$\sum_{1 \le w \le |W|} k_{i,w} = 1 \quad i \in N, \tag{4.2}$$

$$s_{i,w} \ge k_{i,w} \cdot a_i \quad i \in N, w \in W, \tag{4.3}$$

$$s_{i,w} \le M \cdot k_{i,w} \quad i \in N, w \in W, \tag{4.4}$$

$$g_{i,w} \ge k_{i,w} \cdot r_i \quad i \in N, w \in W, \tag{4.5}$$

$$g_{i,w} \le M \cdot k_{i,w} \quad i \in N, w \in W, \tag{4.6}$$

$$f_{i,w} \ge s_{i,w} + g_{i,w} \quad i \in N, w \in W, \tag{4.7}$$

$$W_{max} - M \cdot (k_{i,w} - 1) - w \ge 0 \quad i \in N, w \in W, \tag{4.8}$$

$$\xi_{i,j,w} + \xi_{j,i,w} = 1 \quad i \in N, j \in N, w \in W, \tag{4.9}$$

$$f_{i,w} - s_{j,w} \leq M \cdot (3 - k_{i,w} - k_{j,w} - \xi_{i,j,w}) \\ i \in N, j \in N, w \in W, \tag{4.10}$$

$$d_i \geq f_{i,w} - k_{i,w} \cdot a_i - M \cdot (1 - k_{i,w}) \quad i \in N, w \in W, \tag{4.11}$$

$$d_i \leq t_i \quad i \in N. \tag{4.12}$$

Constraint (2) ensures that one request can only be assigned to a wavelength. Constraints (3) and (4) guarantee that the transmission of a request cannot be started before its arriving time. Constraints (5) and (6) ensure the granted time periods for a request must be greater than its required time periods. Constraint (7) ensures that the transmission of a request completes when all data of the request has been transmitted. Constraint (8) denotes the maximum index of the required wavelengths. Constraints (9) and (10) are nonoverlapping constraints. Constraint (11) implies the delay of a request. constraint (12) ensures that the latency requirements of all requests are satisfied. Moreover, the time complexity of the ILP model is $O(WN^2)$, where W and N are the number of candidate wavelengths and the number of transmission requests, respectively.

As discussed in Section III, two potential architectures are adopted in the FMC access network: the fixed and mobile transmission utilize only their respective wavelengths in the separate architecture or the fixed and mobile transmission share all wavelengths in the sharing architecture. In the separate architecture, we separately input a fixed traffic matrix and a mobile traffic matrix to calculate their respective minimum number of required wavelengths. In this case, their sum is the minimum number of required wavelengths for the network. In the sharing architecture, we input the traffic matrix that is a mixture of the fixed request matrix and the MFH transmission request matrix to calculated the minimum number of required wavelengths.

4.4 Proposed algorithm

For the sharing architecture and the fixed request component of the separate architecture, Problems (1) and (13) are variants of the traditional scheduling problem, which has been shown to be NP-hard [54]; therefore, Problems (1) and (13) are also NP-hard. As the size of the network and the number of traffic requests increases, it becomes difficult to obtain the optimal solution to Problems (1) and (13) in an acceptable time. To comprehensively compare the two potential architectures, we propose an heuristic algorithm for calculating the number of required wavelengths and the transmission latency for the sharing architecture in large-scale networks, which are commonly close to in real-world scenarios.

The objective of the algorithm is to determine the minimal number of required wavelengths in the FMC access network while satisfying the latency requirements of both fixed and mobile transmission. The basic idea of the algorithm is to assign the MFH requests first in each millisecond because the delay requirement of MFH transmission is much stricter than that of fixed transmission. The notations of the algorithm are shown as follows:

Notations:

A : set of request arriving time, where, $a_i \in A$.

R : set of request data size, where, $R_i \in R$.

T : set of request latency requirement, where, $t_i \in T$.

i : index of request i.

j : index of millisecond j.

W_{mobile} : the minimal number of required wavelengths for MFH transmission.

R_m : required time period (determined by the data size) for a TDD burst.

R_i : required time period (determined by the data size) for request i.

N_m : the number of ONUs connected to an RU.

WP : the wait pool.

D_i : the delay of request i.

Fin_k : the transmission completion time of λ_k.

The steps in the algorithm are designed as follows:

Step 1: We first sort A, R and T in increasing order of arrival time a_i. Then, we calculate the the initial value of required wavelengths W_{mobile} for MFH transmission in Formula (14).

$$W_{mobile} = \lceil N_m \lfloor R_m/250 \rfloor \rceil \tag{4.13}$$

Step 2: In each millisecond j, we first assign TDD bursts (i.e., the MFH request). We assign each TDD burst to the wavelength λ_k with the earliest transmission completion time Fin_k, calculate the delay D_i, update λ_k and Fin_k, as listed in Lines 4 to 6. After all TDD bursts arrived in this millisecond j have been transmitted, we traverse the fixed requests in the wait pool and determine whether they can be assigned to the wavelength λ_k with the earliest transmission completion time Fin_k according to Line 9. Then, we calculate the delay D_i, update λ_k and Fin_k, remove the assigned fixed requested from the wait pool.

Step 3: We assign the fixed requests to wavelengths λ_1,λ_2,...,$\lambda_{W_{mobile}}$ in order of their arrival time in millisecond j. We determine whether they can be assigned to

Algorithm 2 Heuristic algorithm

1: for the ith request arrived in jth millisecond
2: **if** $t_i = 250$ **then**
3: Counter $\leftarrow$ Counter + 1
4: Assign request i to the wavelength λ_k with the
5: earliest transmission completion time Fin_k,
6: calculate the delay D_i, update λ_k and Fin_k.
7: **if** Counter mod $N_m == 0$ **then**
8: Traverse the fixed requests R_l in WP
9: **while** $max\{Fin_k, a_l\} + R_l \leq j * 1000 + 1000$ **do**
10: Assign request l to the wavelength λ_k, calculate the
11: delay D_l, update λ_k and Fin_k,
12: remove request l from WP.
13: **end while**
14: **end if**
15: **else**
16: **if** $max\{Fin_k, a_i\} + R_i \leq j * 1000 + 1000$ **then**
17: Assign request i to the wavelength λ_k, calculate the
18: delay D_i, update λ_k and Fin_k,
19: **else**
20: Add request i to WP
21: **end if**
22: **end if**

the wavelength λ_k with the earliest transmission completion time Fin_k according to Line 16. Then, we calculate the delay D_i, update λ_k and Fin_k if request R_i can be assigned, or we add request R_i to the wait pool if request R_i can not be assigned.

If the delay of any fixed requests the latency requirements, we stop the assignment, set $W_{mobile} \leftarrow W_{mobile} + 1$ and recalculate from Step 2. If the delay of all fixed requests is less than 2.5 milliseconds, we output W_{mobile} and the delay D_i for each request i as the result of “Sharing Architecture”.

As for “Separate Architecture”, we can calculate the required wavelengths and the delay by input the matrix of fixed requests and MFH transmission requests separately.

4.5 Numerical analysis

In this section, we analyze the results calculated by the developed ILP model and the proposed algorithm of an FMC access network that is based on tree topology.

We assume that N ONUs connected to fixed users and M ONUs connected to RUs are deployed in the FMC access network. We set different values of M and N to evaluate the impact of varying the number of ONUs connected to fixed users and/or RUs. Considering the overhead (using RS(255, 223) as forward error correction (FEC) leads to approximately 13% overhead [63]), and the upstream bit rate is 8.7 Gbps per wavelength. The experimental parameters are shown in Table. 1.

Table 4.1: Parameters of experiments

Parameter	Value
Delay threshold	250 μs (MFH) / 2.5 ms (fixed)
Upstream link bandwidth	8.7 Gbps per wavelength
Data size of a TDD burst	97896 * 2 * N_{MIMO}
N_{MIMO}	1, 2 and 4
N_{ONU} for fixed users	N
N_{ONU} for MFH	M

Upstream traffic volumes are prioritized since the experiments are conducted in a static scenario. For fixed upstream traffic, we assume that each ONU generates 10 upstream traffic requests as the input. The upstream traffic of the fixed transmission is generated as self-similar traffic, which can be obtained by aggregating several substreams with ON/OFF periods that follow the Pareto distribution [55]. Moreover, we consider the offered network load of the ONUs for fixed services to be the ratio of the fixed traffic requests generated per second and the total capacity of all wavelengths, and we set the network load for fixed services to 0.7 in all default cases. For mobile requests, we consider TDD-based MFH transmission with one TDD burst generated by each RU per millisecond. The maximum data size of the burst is determined by $S_{TB}N_{TB}N_{MIMO}$, where S_{TB}, N_{TB} and N_{MIMO} represent the data size of transport block, the number of transport blocks in each subframe and the number of multiple-input multiple-output (MIMO) layers, respectively. According to [64], we set S_{TB} to 97896 bits, N_{TB} to 2 and N_{MIMO} to 1, 2 and 4, respectively. We assume that each RU generates 10 TDD bursts and all ONUs transmit the TDD blocks with the maximal size to the OLT simultaneously. Moreover, we set the latency thresholds of fixed and MFH upstream transmission to 250 μs and 2.5 ms [39], respectively. The results are evaluated as the average of 50 iterations by using the GUROBI optimizer v9.5.2 [58]. The program is running on a computer with 32 GB of memory and an AMD 6-core, 12-thread, 3.7GHz processor.

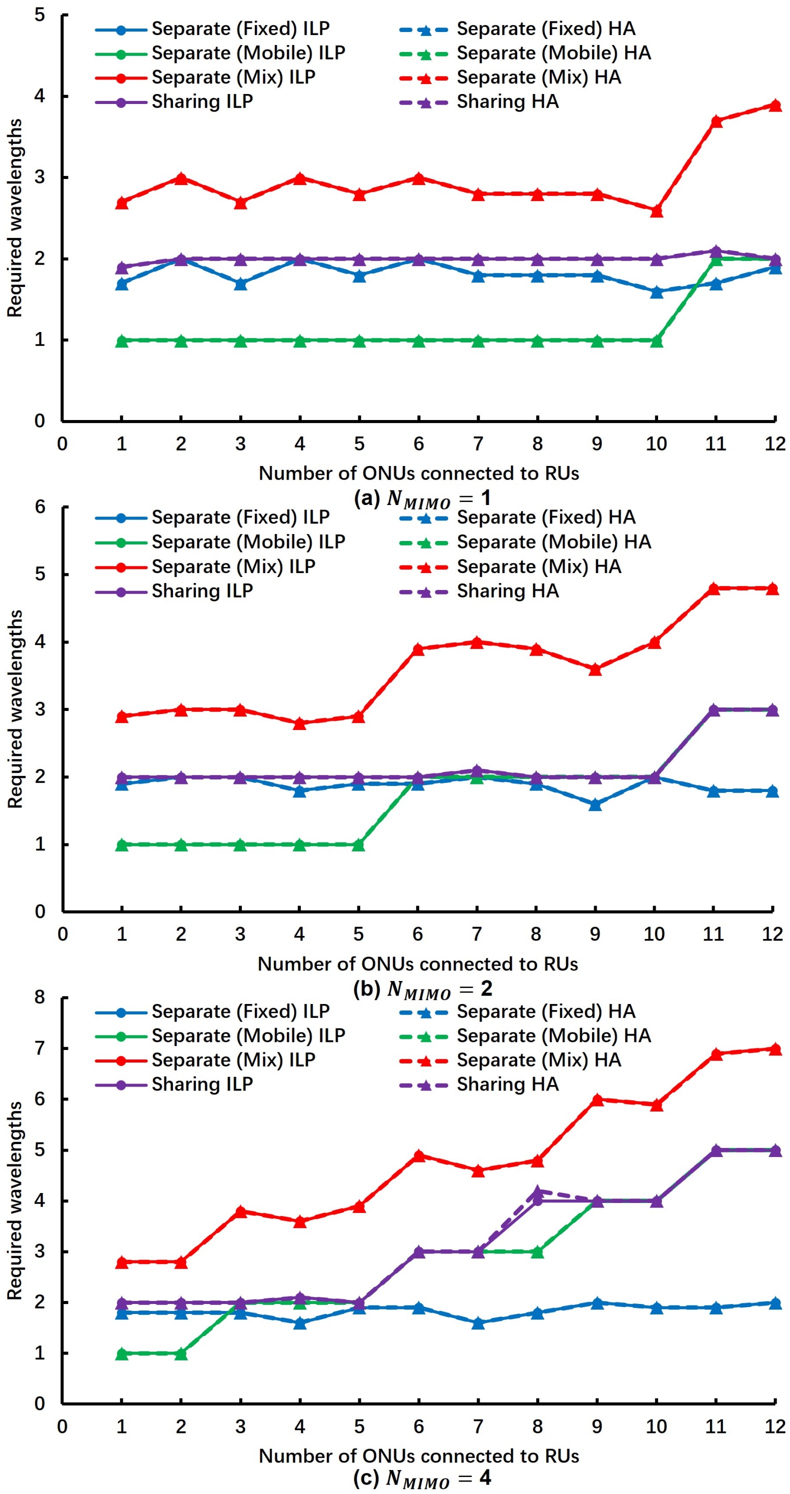

Figure 4.1: The number of wavelengths for the network when different numbers of ONUs are connected to RUs: (a) $N_{MIMO} = 1$; (b) $N_{MIMO} = 2$; (c) $N_{MIMO} = 4$.

For the separate architecture, we separately input the fixed request matrix and the MFH transmission request matrix to calculate the number of required wavelengths. The results are indicated as "Separate (Fixed)" and "Separate (Mobile)", respectively. the number of required wavelengths for the separate architecture is the sum of "Separate (Fixed)" and "Separate (Mobile)", which is denoted as "Separate (Mix)". For the sharing architecture, we input a mixture of the fixed request matrix and the MFH transmission request matrix to calculate the number of required wavelengths that is indicated as "Sharing".

In Fig. 4.1, we show the impact of the number of ONUs connected to RUs and the N_{MIMO} value on the number of required wavelengths. We set N to 8, M is varied from 1 to 12 and N_{MIMO} is set to 1, 2 and 4, respectively. The solid and the dashed lines show the results calculated by the ILP models and heuristic algorithms (HA), respectively. The maximum gap between the solid and dashed lines is 5%, demonstrating that the algorithms can effectively calculate the required number of wavelengths for FMC access networks. Moreover, the sharing architecture requires fewer wavelengths than the separate architecture in all cases. It is because in the sharing architecture, the fixed requests and the MFH transmission requests can share wavelengths, while in the separate architecture, the fixed requests and the MFH transmission requests are assigned to different wavelength domains. In addition, as the value of N_{MIMO} increases (i.e., the maximum data size of a TDD burst increases), the impact on the number of required wavelengths for the MFH transmission caused by the number of ONUs connected to RUs increases. This observation occurs because the number of TDD bursts assigned to a wavelength decreases as the maximum data size of the TDD burst increases. In most cases, the sharing architecture requires that the number of wavelengths be equal to or slightly greater than the larger one required by the fixed or the MFH transmission components in separate architecture. This finding indicates that the fixed requests in the sharing architecture can be assigned to time slots in each millisecond that the TDD bursts cannot be assigned to (the TDD bursts can be assigned to only the first 250 μs in each millisecond because of the latency requirement of MFH transmission). This result confirms that the sharing architecture can provide higher wavelength utilization than the separate architecture.

Fig. 4.2 shows the delay distribution of the MFH transmission calculated by the proposed algorithm corresponding to Fig. 4.1. The red and purple lines represent the separate architecture and the sharing architecture. We can see that the maximum delay of MFH transmission is less than the MFH latency threshold of 250 μs in all cases. This result indicates that the proposed algorithm can calculate the number of required wavelengths for the separate architecture and the sharing architecture while satisfy the latency requirements of the MFH transmission.

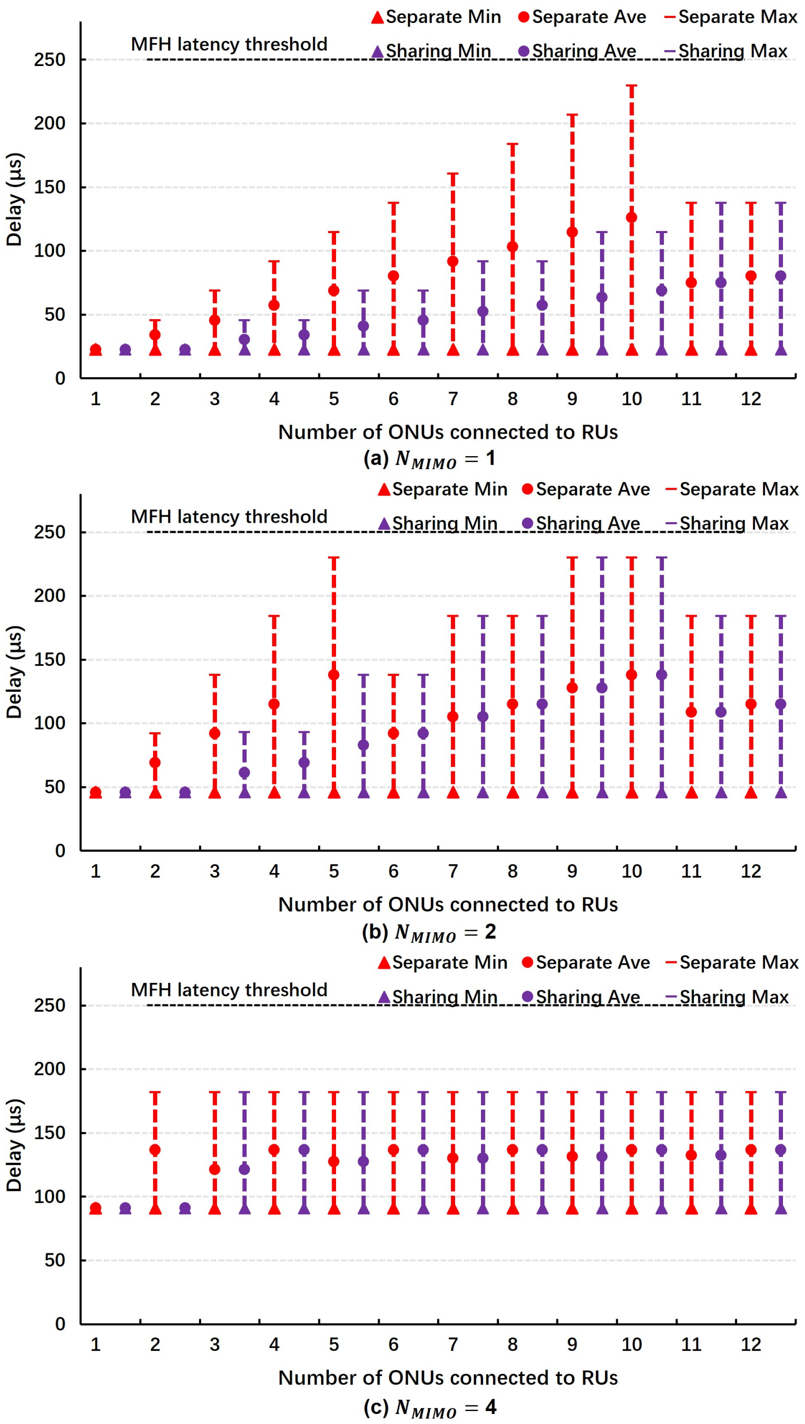

Figure 4.2: Latency distribution when different numbers of ONUs are connected to RUs: (a) $N_{MIMO} = 1$; (b) $N_{MIMO} = 2$; (c) $N_{MIMO} = 4$.

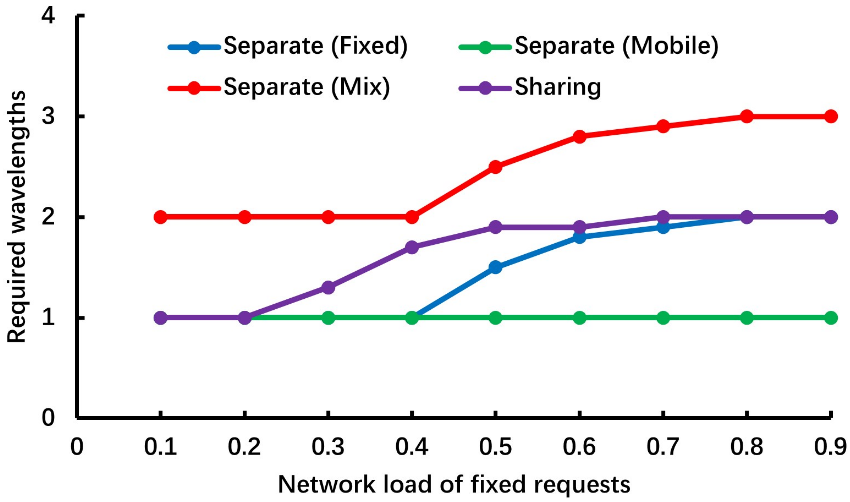

Figure 4.3: The number of wavelengths for the network with different fixed request loads.

Fig. 4.3 shows the impact of the fixed requests load in the FMC access network on the number of required wavelengths. In this case, we set N to 8, M to 8 and N_{MIMO} to 2. As the network load increases, the separate architecture and the sharing architecture both require more wavelengths to satisfy the latency requirements for fixed requests. The sharing architecture requires fewer wavelengths than the separate architecture under all network loads. Moreover, the sharing architecture starts to require more wavelengths when the network load reaches 0.2, which is lighter than the network load (0.4) when the separate architecture starts to require more wavelengths. This result indicates that fixed requests and MFH transmission requests can share the wavelengths in the sharing architecture. Furthermore, when the network load reaches 0.7, the numbers of required wavelengths for the separate architecture starts and the sharing architecture are close to the maximum value. This observation explains why we set the network load to 0.7 in other experiments.

Fig. 4.4 and Fig. 4.5 shows the number of required wavelengths and the delay distribution of MFH transmission in large-scale networks. In this case, we set N set to 32, M is varied from 4 to 48 and N_{MIMO} to 1, 2 and 4, respectively. The results are similar to Fig. 4.1 and Fig. 4.2. Moreover, we can see that the FMC access network is more sensitive to the increase to the number of ONUs connected to RUs. When the number of ONUs increases by 4 times, the number of required wavelengths for the fixed transmission increases by approximately 2.5 times (from approximately 2 in Fig.4.1 to approximately 5 in Fig. 4.4), while the number of required wavelengths for the MFH transmission increases approximately 4 times. It is because the MFH transmission has stricter latency requirement than the fixed transmission.

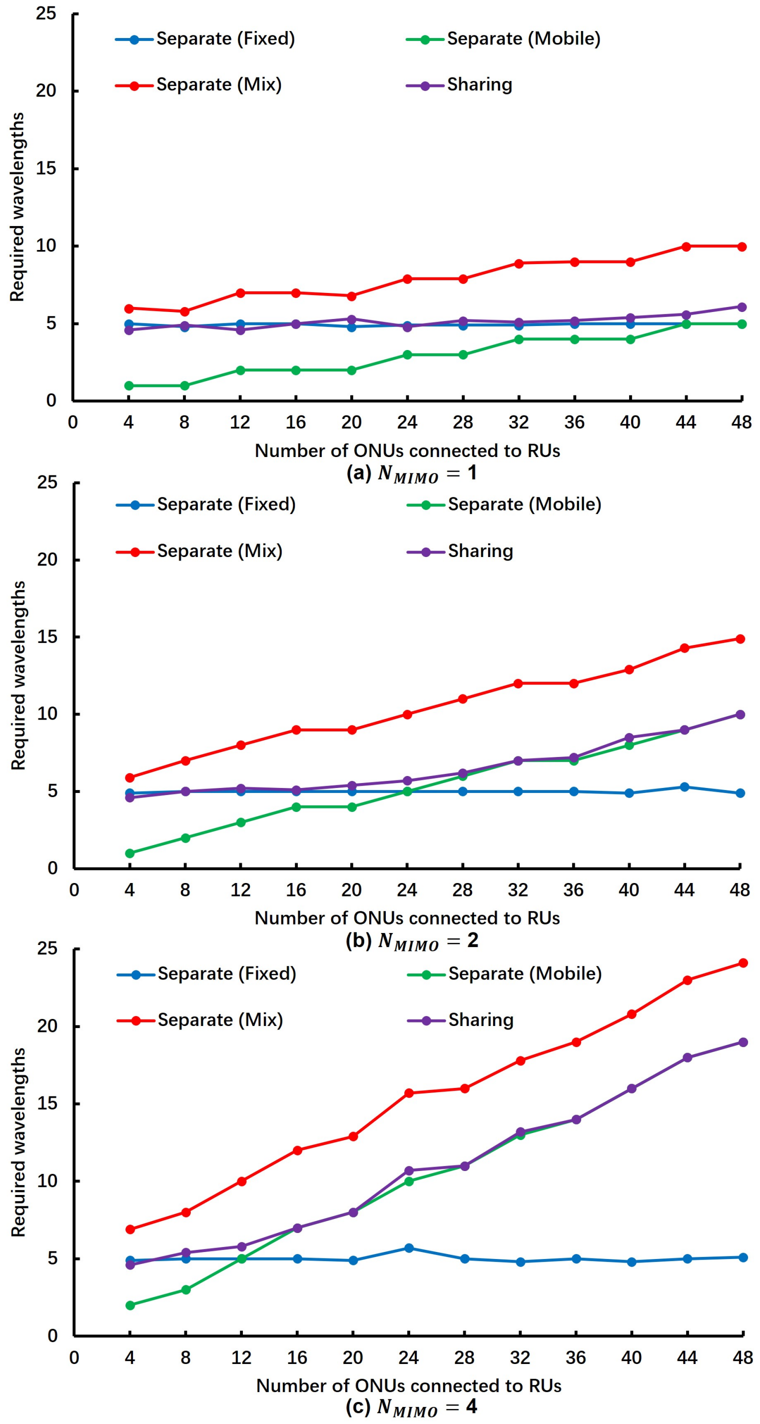

Figure 4.4: The number of wavelengths for the network when different numbers of ONUs are connected to RUs in large-scale networks: (a) $N_{MIMO} = 1$; (b) $N_{MIMO} = 2$; (c) $N_{MIMO} = 4$.

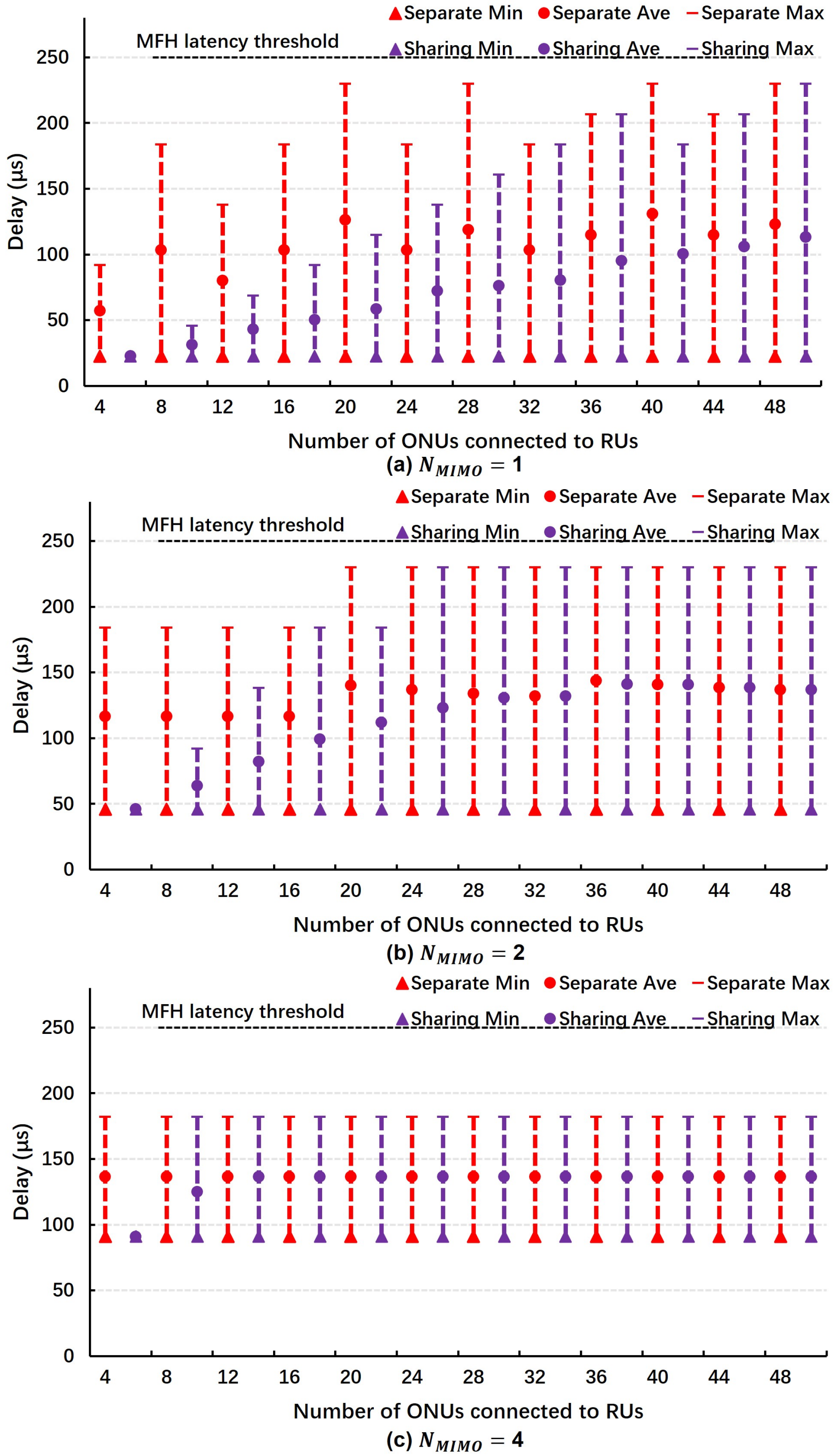

Figure 4.5: Latency distribution when different numbers of ONUs are connected to RUs in large-scale networks: (a) $N_{MIMO} = 1$; (b) $N_{MIMO} = 2$; (c) $N_{MIMO} = 4$.

4.6 Conclusion

In this research, we investigate the wavelength deployment scheme for TWDM-PON-based FMC access networks. We develop an ILP model to minimize the number of required wavelengths in this network. We perform experiments and compared two potential architectures. From the experimental results, the separate architecture and the sharing architecture both satisfy the latency and bandwidth requirements of MFH transmission. The separate architecture is easier to implement since the fixed and the mobile transmission are separated from wavelength domains and fewer wavelengths are required in an ONU, while the sharing architecture provides higher bandwidth utilization to save wavelength resources since the wavelengths are shared by fixed and MFH transmission requests. Network operators can select a suitable architecture before establishing future FMC access networks according to their preferences. The overall motivation of this research is to reduce deployment costs for network operators.

Chapter 5

Summary

In this thesis, we address the design problems of next generation access network based on PONs. Specifically, we discuss the wavelength deployment scheme selection problem in a static scenario in terms of different architectures of PONs and characteristics of traffic demands. We summarize this thesis as follows:

In Chapter 1, we introduce the position of our works, followed by a brief description of the content and contributions of our works. In addition, we show the organization of this thesis.

In Chapter 2, we introduce the background of the PONs including the overview of PONs, architecture and technologies of next generation PONs, as well as the future application of PONs in mobile x-haul transmission and the challenges brought by the potential application.

In Chapter 3, we discuss the design of ONUs and wavelength deployment schemes in an NG-EPON. Specifically, we try to find the most appropriate wavelength deploy scheme by considering the trade-off between the network performance (i,e,. the upstream latency) and the device cost (i,e,. the ONUs cost). We develop an ILP model to solve this problem in a static EPON scenario and propose a heuristic WDSCS algorithm to efficiently solve the problem in a large network scenario. According our experimental results, we have provided a solution for network operators to select an appropriate architecture for next-generation PONs based on user requirements as well as quality of service (QoS) requirements.

In Chapter 4, we investigate the wavelength deployment scheme for TWDM-PON-based FMC access networks. we aim at minimizing the number of wavelengths required for the network while satisfying the latency requirement of both mobile x-haul transmission and fixed transmission. Moreover, we also compare the two potential architectures proposed in previous works. To accomplish this, we develop an ILP model and propose a heuristic algorithm. From our experimental results, we have provided a solution for network operators so that they can select a suitable architecture before establishing future FMC access networks according to their preferences.

Chapter 6

Future work

Future works will mainly focus on the application of PONs in the data center. A data center is the network infrastructure with the purpose of storing, computing, disseminating and presenting data. Data center is critical to an enterprise or organization, so that there are millions of data centers around the world, and almost all enterprise and government have established and maintained their data centers. The architectures of data center networks generally consist of several layers, where the switches in the upper layers are connected to multiple switches in the lower layers. The connection between the switches in adjacent layers is in a P2MP manner; therefore, PONs are considered to be deployed in data center networks to connect adjacent layer, which can significantly save the cost of deployment and maintenance [65]. However, a data center network has its unique requirements different from an access network. On the one hand, a data center network may consist of several PONs cascading with each other to from a hierarchical architecture and data transmissions may occur across various level switches. In some intermediate layers, each node may be connect to all switches in the lower layers, which leads to several PONs connected in parallel in this layer. These features of data center introduce a new and more complicated routing problem that do not exist in a PON-based access network. Therefore, the bandwidth and wavelength assignment problem in a PON-based access network evolves to be the multi-path routing, bandwidth and wavelength assignment problem in a PON-based data center. On the other hand, a major concern of data centers is the power consumption. As there are always thousands of servers in a data center, minimizing the PON power consumption is extremely important. Therefore, the multi-path routing and resource assignment problem must be solved in an energy-efficient approach. In this case, we may need to address the challenges including energy-aware off-loading from switches, avoiding using tunable lasers and so on. Therefore, designing the PON-based data center architecture is a challenging task.

Acknowledgements

First, my heartiest thanks flow to my advisor, professor Yongbing Zhang, for all his guidance and help. His serious scientific attitude, rigorous academic spirit and noble personality have had a great impact on my life. From the selection of the project to the final completion of the project, He has always given me careful guidance, which not only set me up for ambitious academic goals, but also helps me a lot to do my further research, and at the same time is still in the spirit and life It gave me meticulous care.

With many thanks to Mingcong Yang, the Senior Engineer in Huawei Wuhan Research Center, Huawei, China, for his for his inspiration and advice, which helped me to significantly improve my researches.

I would like to thank my friends and classmates, for their companion in my daily life in the past three years.

I want to express my sincere gratitude to the JST Student Research Fellowship Support Plan - Support for Pioneering Research Initiated by the Next Generation (SPRING) for the financial support, which helped me me to focus on my research and complete my Ph.D. course without financial pressure.

Finally, from the bottom of my heart I would like to thank my parents, for their love, understanding, support and comfort during these years, which are the key pillars to my Ph.D. degree.

Bibliography

[1] K. Kumozaki, "Optical access systems: present state and future directions," *NTT Tech. Rev*, vol. 6, pp. 1-7, 2008.

[2] Cisco White Paper, Online, "Cisco Annual Internet Report (2018–2023) White Paper," Available: https://www.cisco.com

[3] S. Arnon, J. Barry, G. Karagiannidis, R. Schober, and M. Uysal, *Advanced optical wireless communication systems.* Cambridge university press, 2012.

[4] G. Pesavento and M. Kelsey, "PONs for the broadband local loop," *Lightwave, PennWell.*, vol. 16, no. 10, pp. 68-74, Sept. 1999.

[5] B.Lung, "PON architecture 'futureproofs' FTTH," *Lightwave, PennWell.*, vol. 16, no. 10, pp. 104-107, Sept. 1999.

[6] C. Zhang et al., "Analysis of wavelength deployment schemes in terms of optical network unit cost and upstream transmission performance in NG-EPONs," *J. Opt. Commun. Netw.*, vol. 13, no. 9, pp. 214-223, Sept. 2021.

[7] J. R. Stern, J. W. Balance, D. W. Faulkner, S. Hornung, D. B. Payne, and K. Oakely,"Passive optical local networks for telephony applications and beyond,"*Electron. Lett.*, vol. 23, no. 24, pp. 1255-1256, 1987.

[8] FSAN, Online, Available: https://www.fsan.org/.

[9] ITU-T Recommendation G.983 Series, "Broadband passive optical network," 2001.

[10] ITU-T Recommendation G.984 Series, "Gigabit capable passive optical network (G-PON)," 2008.

[11] IEEE 802.3ah-2004. "IEEE standard for information technology–Local and metropolitan area networks–Part 3: CSMA/CD access method and physical layer specifications amendment: Media access control parameters, physical layers, and management parameters," IEEE Std 2004.

[12] IEEE 802.3av-2009,"IEEE standard for information technology, Local and metropolitan area networks, Specific requirements, Part 3: CSMA/CD access method and physical layer specifications amendment 1: Physical layer specifications and management parameters for 10 Gbps passive optical networks," 2009.

[13] ITU-T Recommendation G.9807 Series,"10 Gigabit capable symmetric passive optical network (XGSPON)," 2016.

[14] ITU-T Recommendation G.987 Series, "10 Gigabit capable passive optical network (XG-PON)," 2012.

[15] ITU-T Recommendation G.989.1, "40-Gigabit-capable passive optical networks (NG-PON2): general requirements," 2013.

[16] IEEE 802.3 Ethernet Working Group, "IEEE P802.3ca 50G-EPON Task Force," Online, Available: http://www.ieee802.org/3/ca/.

[17] C. Knittle, "IEEE 100G-EPON,"," *Optical Fiber Communication Conf.*, 2016.

[18] G. Kramer and G. Pesavento,"Ethernet passive optical network (EPON): building a next-generation optical access network," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 66-73, Aug. 2002.

[19] Ed Harstead, D. van Veen, V. Houtsma and P. Dom,"Technology Roadmap for Time Division Multiplexed Passive Optical Networks(TDM PONs)," *J. Lightwave Technol.*, vol. 37, no. 2, pp. 657-664, 2018.

[20] G. Kramer,"Ethernet Passive Optical Networks," 2005.

[21] G. Kramer, B. Mukherjee and G. Pesavento,"IPACT a dynamic protocol for an Ethernet PON (EPON)," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 74-80, Aug. 2002.

[22] Houtsma, V. Mahadevan, A. Kaneda, N. Veen and Doutje "Transceiver technologies for passive optical networks: past, present, and future," *J. Opt. Commun. Netw.*, vol. 13, no. 1, pp. A44-A55, 2021.

[23] P. P. Iannone, et al., "Increasing TDM rates for access systems beyond NG-PON2," *J. Lightwave Technol.*, vol. 34, pp. 1545–1550, 2016.

[24] Y. Luo et al., "Time- and Wavelength-Division Multiplexed Passive Optical Network (TWDM-PON) for Next-Generation PON Stage 2 (NG-PON2)," *J. Lightwave Technol.*, vol. 31, no. 4, pp. 587-593, 2013.

[25] V.Houtsma, D. van Veen and Ed.Harstead,"Recent progress on standardization of next-generation 25, 50, and 100G EPON," *J. Lightwave Technol.*, vol. 35, no. 6, pp. 1228-1234, Aug. 2017.

[26] D. van Veen et al., "High Speed TDM-PON Beyond 10G" *Optical Fiber Communications Conf. and Exhibition (OFC)*, 2016

[27] V.Houtsma, D. van Veen and Ed.Harstead, "Recent Progress on 25G EPON and beyond," *European Conference on Optical Communication*, pp. 1-3, Sept. 2016.

[28] G. Kramer, "A proposal for channel bonding at MAC control sublayer," *IEEE 802.3ca 100G-EPON Task Force*, 2016.

[29] F. Effenberger et al., "MPCP+ Issues and solutions," *HUAWEI TECHNOLOGIES CO.*, 2016

[30] L. Zhang, Y. Luo, B. Gao, X. Liu, F. Effenberger, and N. Ansari, "Channel bonding design for 100 Gb/s PON based on FEC code word alignment," *Optical Fiber Communications Conf. and Exhibition (OFC)*, 2017.

[31] Cisco, "Global_2021_Forecast_Highlights," Online, Available: https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2021_Forecast_Highlights.pdf.

[32] FTTH Council Europe, Online, https://www.ftthcouncil.eu/.

[33] A. Pizzinat, P. Chanclou, T. Diallo, and F. Saliou, "Things you should know about fronthaul," *J. Lightw. Technol.*, vol. 33, no. 5, pp. 1077–1083, Mar. 2015.

[34] M. Jaber et al., "A joint backhaul and RAN perspective on the benefits of centralised RAN functions," *IEEE International Conference on Communications Workshops (ICC)*, 2016.

[35] Common Public Radio Interface (CPRI), Online, Available: http://www.cpri.info

[36] European Telecommunications Standards Institute, Group Specification, Open Radio equipment Interface (ETSI GS ORI) 001, Requirements for ORI (Release 4), 2014.

[37] A. Jaber et al., "A survey of 5G network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[38] Common Public Radio Interface (CPRI); eCPRI Interface Specificatio V1.0, Aug. 2017.

[39] Study on New Radio Access Technology: Radio Access Architecture and Interfaces, document 3GPP TR 38.801, Mar. 2016.

[40] O-RAN Fronthaul Working Group, "Control, User and Synchronization Plane Specification", ORAN.WG4.CUS.0-v03.00, 2020.

[41] R. Esmailzadeh, et al., "Prerake diversity combining in time-division duplex CDMA mobile communications," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 3, pp. 795-801, 1999.

[42] K. Pedersen, et al., "A flexible 5G frame structure design for frequency-division duplex cases," *IEEE Commun. Mag.*, vol. 54, no. 3, pp. 53-59, 2016.

[43] Huo. Y, Dong. X and Xu. W,"5G cellular user equipment: From theory to practical hardware design," *IEEE Access.*, vol. 5, pp. 13992-14010, 2017.

[44] J.Kani, J.Terada, K.Suzuki and A. Otaka, "Solutions for future mobile fronthaul and access-network convergence," *J. Lightwave Technol.*, vol. 35, no. 3, pp. 527-534, 2016.

[45] R. Muñoz, R. Vilalta, R. Casellas, R. Martinez, R. Szyrkowiec, T. Autenrieth, A. López, V. López and Diego, "Integrated SDN/NFV management and orchestration architecture for dynamic deployment of virtual SDN control instances for virtual tenant networks," *J. Opt. Commun. Netw.*, vol. 7, no. 11, pp. B62-B70, 2015.

[46] X. Liu and E. Frank, "Emerging optical access network technologies for 5G wireless," *J. Opt. Commun. Netw.*, vol. 8, no. 11, pp. B70-B79, 2016.

[47] E. Wong, M. Mueller, M. P. Dias, C. A. Chan, and M. C. Amann,"nergy-efficiency of optical network units with vertical-cavity surface-emitting lasers," *Opt. express*, vol. 20, no. 14, pp. 14960-14970, 2012.

[48] A. Helmy, H. Fathallah, and H. Mouftah,"Interleaved Polling Versus Multi-Thread Polling for Bandwidth Allocation in Long-Reach PONs," *J. Opt. Commun. Netw.*, vol. 4, no. 3, pp. 210-218, Mar. 2012.

[49] L. Wang, X. Wang, M. Tornatore, H. S. Chung, H. H. Lee, S. Park, and B. Mukherjee,"Dynamic Bandwidth and Wavelength Allocation Scheme for Next-Generation Wavelength-Agile EPON," *J. Opt. Commun. Netw.*, vol. 9, no. 3, pp. B33-B42, 2017.

[50] S. B. Hussain, W. Hu and H. Xin and A. M. Mikaeil,"Low-latency dynamic wavelength and bandwidth allocation algorithm for NG-EPON," *J. Opt. Commun. Netw.*, vol. 9, no. 12, pp. 1108-1115, Dec. 2017.

[51] S. B. Hussain, W. Hu and H. Xin, A. M. Mikaeil and A. Sultan,"Flexible Wavelength and Dynamic Bandwidth Allocation for NG-EPONs," *J. Opt. Commun. Netw.*, vol. 10, no. 6, pp. 643-652, 2018.

[52] W. Wang, and W. Guo and W. Hu,"Dynamic Wavelength and Bandwidth Allocation Algorithms for Mitigating Frame Reordering in NG-EPON," *J. Opt. Commun. Netw.*, vol. 10, no. 3, pp. 220-228, Mar. 2018.

[53] M.P. McGarry, M. Reisslein and M. Maier,"WDM Ethernet passive optical networks," *IEEE Commun. Mag.*, vol. 44, no. 2, pp. 15-22, Feb. 2006.

[54] M. Pinedo, Scheduling: Theory, Algorithms, and Systems, Prentice Hall, 2nd ed., 2002.

[55] M.S. Taqqu, W. Willinger and R. Sherman,"Proof of a fundamental result in self-similar traffic modeling," *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 2, pp. 5-23, Mar. 1997.

[56] K. Kanonakis and I. Tomkos,"Improving the efficiency of online upstream scheduling and wavelength assignment in hybrid WDM/TDMA EPON networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 6, pp. 838-848, 2010.

[57] G. Kramer, Online, "Generator of Self-Similar Traffic(version 3)," Available: http://research.glenkramer.com/code/trf_gen3.shtml

[58] Online, "Gurobi v9.1.1." Available: http://www.gurobi.com, Nov. 2020.

[59] T. Tashiro et al., "A novel DBA scheme for TDM-PON based mobile fronthaul," *Optical Fiber Communication Conf.*, 2014. pp. 1-3, Paper Tu3F-3.

[60] T. Kobayashi, H. Ou, D. Hisano, T. Shimada, J. Terada, and A. Otaka, "Bandwidth allocation scheme based on simple statistical traffic analysis for TDM-PON based mobile fronthaul," *Optical Fiber Communication Conf.*, 2016. pp. 1-3, Paper W3C-7.

[61] Y. Nakayama et al., "Efficient DWBA algorithm for TWDM-PON with mobile fronthaul in 5G networks," *IEEE Global Telecommun. Conf.*, Dec. 2017, pp. 1-6.

[62] Y. Nakayama and D. Hisano. "Wavelength and Bandwidth Allocation for Mobile Fronthaul in TWDM-PON," *IEEE Trans. ON Commun.*, vol. 67, no. 11, pp. 7642-7655, 2019.

[63] Google Fiber, Online, "Super-PON PCS Options,"
Available: https://www.ieee802.org/3/cs/public/201907/20190717-PCS_v1.pdf

[64] E-UTRA Physical Layer Procedures, document 3GPP TR 36.213 Release 14.4.0, Sep. 2017.

[65] S. H. Mohamed, et al., "Energy efficiency of server-centric PON data center architecture for fog computing," *IEEE International Conference on Transparent Optical Networks (ICTON)*, pp. 1-4, 2018.