Adaptive Router Bypass Techniques to Enhance Core Network Efficiency

by

Fahad A. Ghonaim

B. Sc., King Abdul-Aziz University, KSA, 2007
M. Eng., University of Victoria, Canada, 2011

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

DOCTOR of PHILOSOPHY

in the Department of Electrical and Computer Engineering

Fahad A. Ghonaim ©

Department of Electrical and Computer Engineering

University of Victoria

March 2018

Adaptive Router Bypassing Techniques to Enhance Network Efficiency

by

Fahad A. Ghonaim

B. Sc., King Abdul-Aziz University, KSA, 2007
M. Eng., University of Victoria, Canada, 2011

Supervisory Committee

Prof. Thomas E. Darcie (Co-supervisor)
(Department of Electrical and Computer Engineering)

Dr. Sudhakar Ganti (Co-supervisor)
(Department of Computer Science)

Dr. Stephen W. Neville (Department member)
(Department of Electrical and Computer Engineering)

Abstract

Internet traffic is increasing exponentially, driven by new technologies such as Internet of Things (IoT) and rich streaming media. The traditional IP router becomes a bottleneck for further Internet expansion due to its high power consumption and inefficiency in processing the growing traffic. Router bypass has been introduced to overcome capacity limitations and the processing costs of IP routers. With router bypass, a portion of traffic is provisioned to bypass the router and is switched by the transport layer. Router bypass has shown to provide significant savings in network costs. These advantages are limited by a reduction in the statistical multiplexing associated with the subdivision of the available bandwidth typically into bypass and traditional portions thus limiting the interest in bypass techniques.

This thesis will explore multiple techniques to enhance the efficiency of router bypass. The main goals are to address the issue of the reduction in statistical multiplexing and to add a dynamic approach to the router bypass mechanism.

The recent advancements in the Optical Transport Network (OTN) play a major role in the transport network. This proposal takes full advantage of OTN in the router-bypassing context by applying recent developments such as Hitless Adjustments ODUflex (HAO), which allow the provisioned channels to be adjusted without re-establishing the connections. In addition, it will allow the bypassing mechanism to be flexible enough to meet the traffic behaviour needs of the future. This thesis will study multiple approaches to enhance the router bypass mechanism including: an adaptive provisioning style using various degrees of provisioning granularities and controlling the provisioning based on traffic behaviour. In addition, this thesis will explore the impact of automation in Software-Defined Networking (SDN) on router

bypass. The application-driven infrastructure in SDN is moving the network to be more adaptive, which paves the way for an enhanced implementation of router bypass.

Many challenges still face the industry to fully integrate the three layers (3, 2, and 1) to transform the current infrastructure into an adaptive application driven network. The IP router (layer 3) provisions and restores the connection regardless of the underlying layers (layer 2 and 1) and the transport layer does the same regardless of the IP layer. Although allowing every layer to develop without being constrained by other layers offers a huge advantage, it renders the transport layer static and not fully aware of the traffic behaviour.

It is my hope that this thesis is a step forward in transforming the current network into a dynamic, efficient and responsive network. A simulation has been built to imitate the router bypassing concept and then many measurements have been recorded.

Contents

List of Figures

List of Tables

List of Equations

Acknowledgments

I would like to thank my co-supervisors: Dr. Darcie and Dr. Ganti. They have helped me immensely in this journey and I would like to express my gratitude to them.

I have worked with Prof. Darcie for the last eight years; he has been a mentor and teacher whom I will always look up to as inspirational thinker. He taught me how to be a true researcher, to think critically and maneuver through complicated problems and find original solutions. He encouraged me to believe that there will always be a solution to whatever we want to achieve.

Dr. Ganti helped me to have positive attitude toward the obstacles that I have faced in my research. He pointed me in the right direction during this journey and I appreciate his patience and the help he offered me as I became a better researcher.

I can not forget my family and parents who have supported me every step of the way.

Dedication

To my parents, family and to every precious soul who stood by me in this journey.

Chapter 1

Introduction

The evolution of Internet content in the last few decades has been dramatic in quantity and quality (e.g., [1]). The Internet advanced from a network intended to transfer simple data (texts and emails) between users to supporting rich media and streaming services. With the revolution of Internet content, the transport network had to evolve to meet the content demand.

1.1 Introduction

Internet traffic growth has been rising steeply for the last decade and prognostications for the future suggest even more dramatic growth. For instance, the projected annual IP traffic for 2018 will be greater than all IP traffic that has been generated globally from 1984 to 2013 [1]. IP traffic will increase threefold in the next five years and the number of connected devices will be three times the global population in 2019 according to forecasts [2]. The Internet of Things (IoT), big data, and rich media streaming are examples of high bandwidth-consuming technologies, which are driving higher-capacity network infrastructure. The wide use of cyper-physical systems (CPS) will add even more to Internet traffic. Moreover, cost reduction in new data-centre deployments with higher demand for cloud services have led to many data centres distributed around the core network [3].

IP-based packet-switching has been a powerful enabler for Internet growth. It allows the statistical sharing of a common transport resource among large

numbers of simultaneous, transient, and diverse connections, all without complex preconnection setup. This sharing provides statistical multiplexing that maximizes bandwidth use, especially in an unpredictable traffic environment; however, IP routing requires considerable per-packet processing which leads to high power consumption. Traditionally, every packet must pass through each intermediate node. It has been shown [4] that, on average, a packet needs to be processed by between 10-15 nodes before it reaches its destination. In core routers, 50-85% of traffic processing resources are consumed by transit traffic [5]. This is compounded by the need to overprovision bandwidth to support peak demand. Busy-hour traffic volume has been reported [6] to exceed average volume by 72% [1].

In contrast, the optical transport layer that provides the capacity for the overlay router networks is highly efficient on a per bit basis. Recent enhancements to the optical transport layer, particularly those introduced in the G.709 optical transport network (OTN) specification, increase its flexibility and ability to interwork with packet networks. Ongoing research seeks to leverage these enhancements and the increased efficiency offered to relieve difficulties in meeting growth demand through enhancements to router networks. This has led many to believe that cross-layer switching will enhance the transport network to be more dynamic and efficient. That efficiency and agility will play an important role in Internet evolution [7].

Router bypass has been introduced as a solution to using optical layer efficiency to enhance traffic transportation in traditional networks. A bypass path is established directly between two network nodes separated by multiple (transit) routers. Traffic is then routed directly through an optical path between these bypass nodes, avoiding the intermediate transit routers. Router bypass has been shown to offer savings in power consumption of up to 45% [8] as well as a reduction in capital and operational expenses of up to 80% and 60%, respectively [5] [9]. However, implementation of the bypass requires the partitioning of network bandwidth into separate bypass and non-bypass portions, and since each portion is smaller than the original total bandwidth, the advantages derived from statistical multiplexing are reduced, resulting in

criticism of router bypassing.

As global IP traffic continues to increase dramatically, driven primarily by large file and streaming transactions [1], network operators seek new methods to increase the efficiency of IP router networks. Given high router complexity and power consumption, traditional IP networking techniques alone struggle to keep pace with bandwidth growth. The statistical multiplexing offered by routers is essential to maximize bandwidth utilization, especially in an unpredictable traffic environment. In contrast, it is well established that the optical layer offers a lower per bit cost and higher capacity transport due to its less complex circuit-switching. Features recently introduced in the G.709 Optical Transport Network (OTN) specification, including direct support mechanisms for packet switching (i.e., Generic Framing Procedure) and the Hitless Adjustment (HAO), paved the way for a more dynamically provisioned transport layer. Yet the optical network is still mainly used for static point-to-point connections while routers manage packet routing. This causes some researchers to believe that a more dynamic transport layer is within reach [7] and this is an attractive alternative for meeting Internet scaling challenges.

1.2 Impact of Internet Traffic Growth

The advancements in electronics manufacturing have had a major impact on the telecommunication industry. For instance, the energy required per bit has been reduced year by year. Figure 1.1-a shows the growing capacity of network equipment as well as the efficiency improvements in electronics on a low level and chips manufacturing as illustrated in Figure 1.1-b. However, the demand for Internet bandwidth in the last decade has grown substantially, pushing the network capacity expansion even further. Predictions suggest that it will keep growing for many more years (see Figure 1.2).
Media consumption, video-streaming applications (such as Netflix, Hulu, YouTube, etc.) as well as data consumed by mobile devices and smart phones are some examples of what have been driving the Internet growth in recent years as shown in Figure 1.2. Recent trends indicate that the annual growth rate of Internet traffic is 50% [10]. According to the Cisco visual networking index,

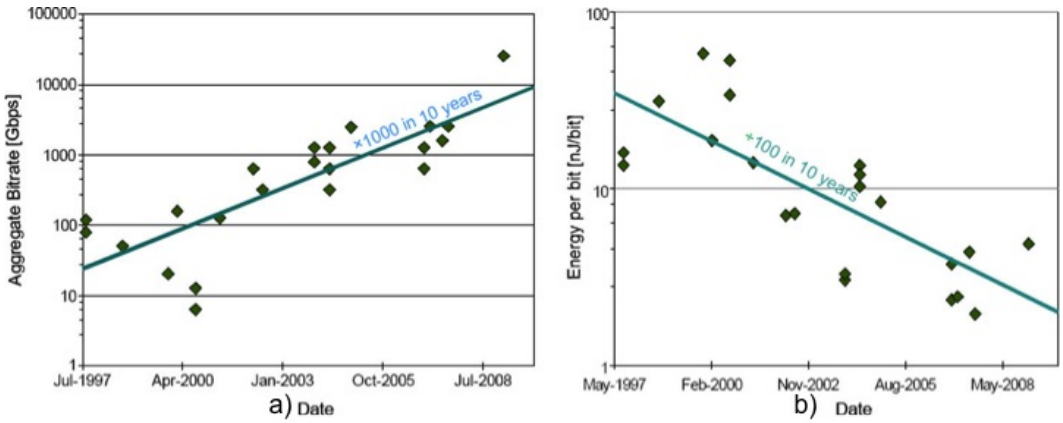


Figure 1.1: Evolution of the bandwidth capacity and energy per bit consumption of last decade [11].

global Internet traffic is expected to increase by threefold in the five years afterwards, following a fourfold growth in five years after that. At some point in 2015, Cisco predicted that global IP traffic would hit 1.0 zettabytes per year, driven by the worldwide growth [12]. The projected annual IP traffic for 2018 will be greater than all IP traffic that has been generated globally from 1984 to 2013 [1] [2]. Moreover, cost reduction in new data-centre deployments combined with higher demand for cloud services have led to a large number of data centres being distributed around the core network [3].

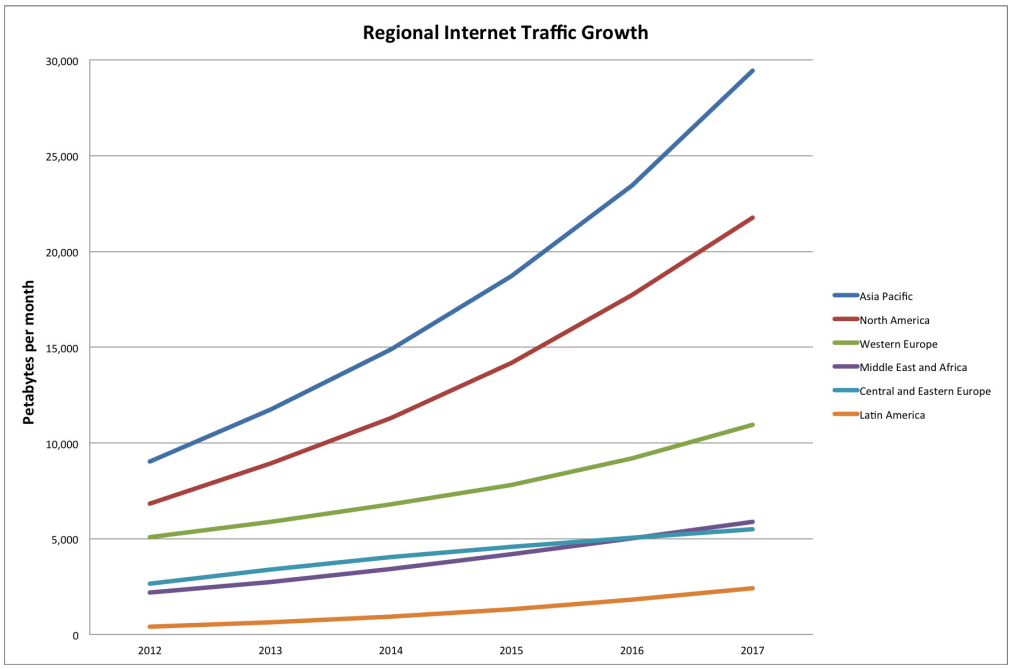


Figure 1.2: Expected Internet traffic growth around the world per month [16]

At the same time, this growth comes with a price, which is tremendous power consumption in the Internet infrastructure. A study shows that the Internet network electricity consumption will grow at rate of 10% per year in the coming years [13].

As we can see in Figure 1.3, the electricity consumption of telecom operators worldwide has risen in the last decade. To get a sense of how much power the core networks consume, a prediction was made by another study [14] that shows that the total electricity consumption of communication networks has exceeded 350TW/h/y in 2013. The power consumption of the Internet in broadband countries is a small percentage (between 1-2%) of the total national power consumption. Considering the cost of one kW is approximately seventeen cents, then cutting only 10% of telecom operator consumption can save up about US$6 billion (assuming the average network consumption of any given country is about 3 TW [8]). If we consider the worldwide consumption mentioned in Figure 1.3 then the savings will be tremendous. It will lead to direct cost savings associated with consumed power and will have major effect on the environment by reducing the carbon footprint of power-hungry equipment. The following is a simple calculation to demonstrate the current impact on the environment: if we consider worldwide electricity consumption in telecom operator networks in 2012 to be around 270 TWh/y, it is equivalent to the annual greenhouse gas emissions of 39.7 millions passenger vehicles [15].

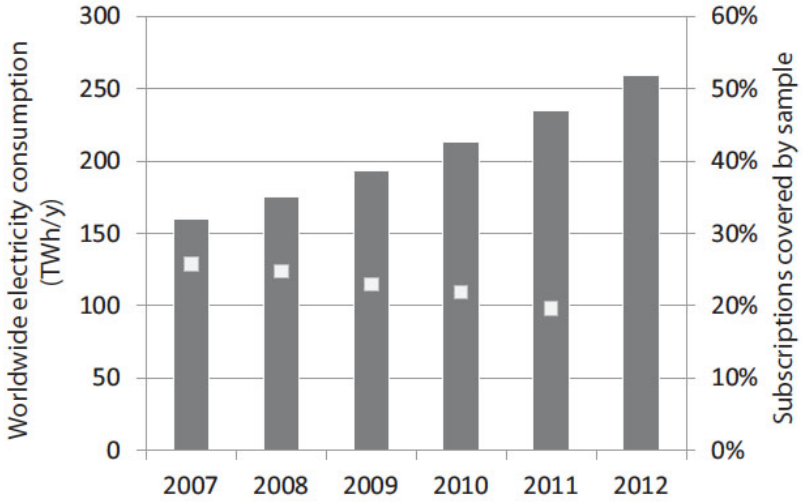


Figure 1.3: Worldwide electricity consumption in telecom operator networks [13].

1.3 IP Routers and the Power Consumption Issue

To find the main source of this power-consumption problem, we need to look at service provider networks and understand which area is consuming the most power. It is clear that routers are one of the main causes of that consumption. For example, if we take an IP over the WDM network, we find that routers consume 90% of the total network power consumption while transponders and EDFAs consume much less, about 7% and 2% of the total power, respectively [8]. To put that in perspective, Cisco CRS-1 16-slot single-shelf system core router consumes approximately 10 kW (when fully configured with line cards in traffic running condition and its switching capacity is 1.2 Tbps) [17].

By taking a closer look at traditional routers, we find that packets processing is responsible for most of the power consumption because the packet size is relatively small compared with the volume of traffic. The maximum packet size of the normal Ethernet is about 1500 bytes but new applications such as VoLTE, P2P, mobile video streaming and online gaming has pushed most of the traffic packet size to under 400 bytes [18]. As mentioned, switching is the costliest process inside the router, requiring that each packet needs to be de-

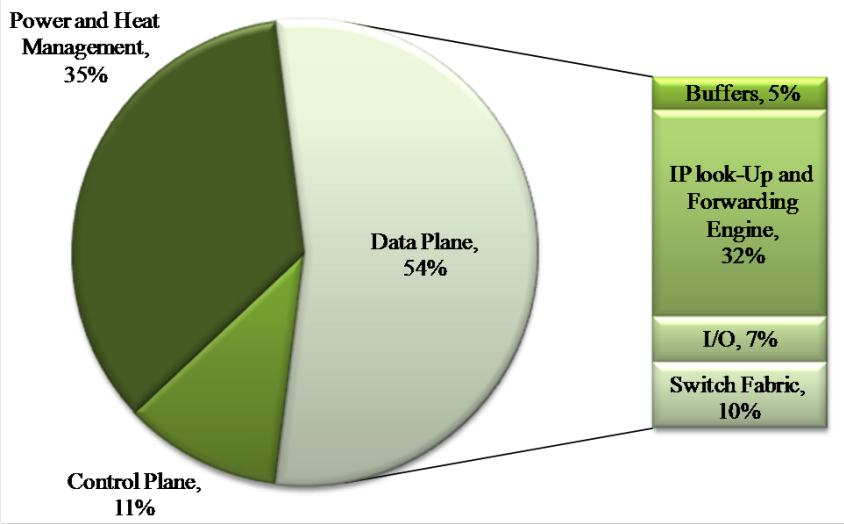


Figure 1.4: Breakdown of the power consumption inside a high end IP router [20].

framed/ framed, buffered and then forwarded. The forwarding part requires IP lookup (IP classification) inside the forwarding engines, which consumes considerable amount of the total power consumption inside the router. For example, the power consumption of IPv4 and IPv6 packet classification with a mere 32 channels at 40 Gb/s can be as high as 700 and 1400 W, respectively [19].

Figure 1.4 shows the breakdown of power consumption inside a high-end router. As a percentage of the total power consumption, we can see that the data plane, which is responsible for processing and switching packets, is consuming about 54%. This includes many elements of the switching system such as forwarding engines, switching fabrics and buffering; therefore, the switching plane alone, in addition to the powering and cooling requirements of the system, consumes about 89% of the total power consumption. Consequently, we note that the control plane—the brain of the router—contains all the routing protocols, routing addresses and operating systems but consumes only about 11%. The conclusion is that the power consumption problem is caused by the traffic being handled/switched within the network infrastructure.

Proposals have been made about optical router architectures [19], because of the efficient nature of optical components in terms of power: however, they

would not be successful in an actual implementation because of the wide gap between electronics and optics in active computer components such as buffers and memory. In this thesis, we will focus on the middle ground and to find a feasible solution to be implemented in the carrier network (since it is based on current technologies) to improve the power consumption issue and meet the demand for bandwidth. This solution is an enhancement of router off-loading or router bypassing.

The advancement of the OTN and sub-lambda switching will open doors to the next generation of optical switching, even when considering higher bandwidth demands. Ultimately, routers will be responsible for routing and addressing (control plane), while switching will be done by the optical network which has higher bandwidth components and is more power efficient than its current alternative. This proposal considers the latest advancement in the optical transport network (OTN) as a base for router bypassing.

1.4 Bandwidth Expansion

In terms of traffic demography, many studies show that most Internet traffic is not uniformly distributed over a network. For example, the top 1% of global subscribers generated more than 20% of all traffic. In addition, the top 10% of global subscribers generated more than 60% of all traffic [12]. From a usage point of view, Internet traffic volume is not consistent over the day. The peak hour has almost 25% more traffic than the average [12]. Network designers cannot expand the network based on peak-hour traffic because it is simply not economically practicable.

Although the Internet content is exploding and evolving, the network infrastructure has not kept up. Many improvements have been made to extend the bandwidth and enhance the overall utilization; however, traffic demography has not been considered in the transport layer. IP router (Layer 3) has its own connection provisioning and restoration mechanisms regardless of the underlying transport layer. At the same time, the OTN or WDM network (transport layer) has its own ways of provisioning and restoration. That sep-

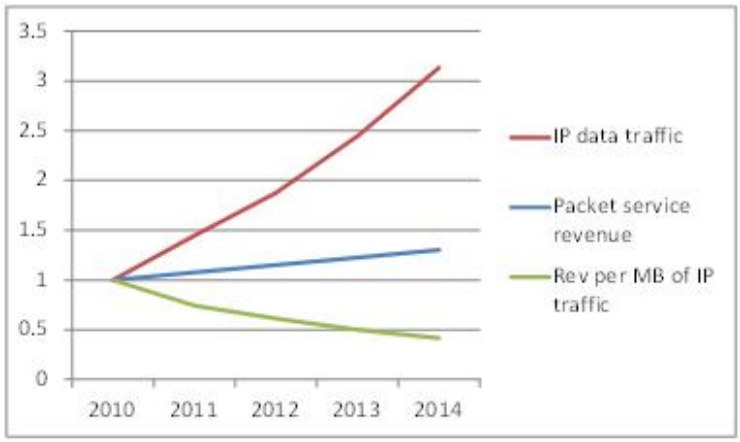


Figure 1.5: Declining revenue per MB for IP traffic [5].

aration has served the network infrastructure in the past where each layer can be developed or replaced without affecting other layers. For instance, the whole IP/SONET layer has been replaced by an OTN layer without any major changes needed in the IP layer.

Although it has had its benefits, the separation between the IP layer and the transport layer has caused the transport layer to become static and unresponsive to the traffic changes. Traffic volume is not only concentrated at certain parts of the network but also, the volume of traffic changes throughout a given day. The changes in traffic behaviour have not been accommodated in the transport layer. That lack of adaptation has led to a lack of efficiency in channel provisioning and traffic switching. For instance, channels are either over provisioned or under provisioned. Figure 1.5 demonstrates a drastic decline in revenues per Mbps for the transit network over the last ten years. That requires more innovations to increase the network efficiency.

We can clearly see the problems service providers are facing in trying to meet the demand for bandwidth. Increasing bandwidth demand, declining revenues, and the rise of power consumption for the core network are issues that show the necessity for the infrastructure of the network to evolve.

Because of the resource-intensive switching process, expanding the capacity of routers has been a challenge. Even with advancement in semiconductors manufacturing, their limited capacity is one of the main obstacles facing the

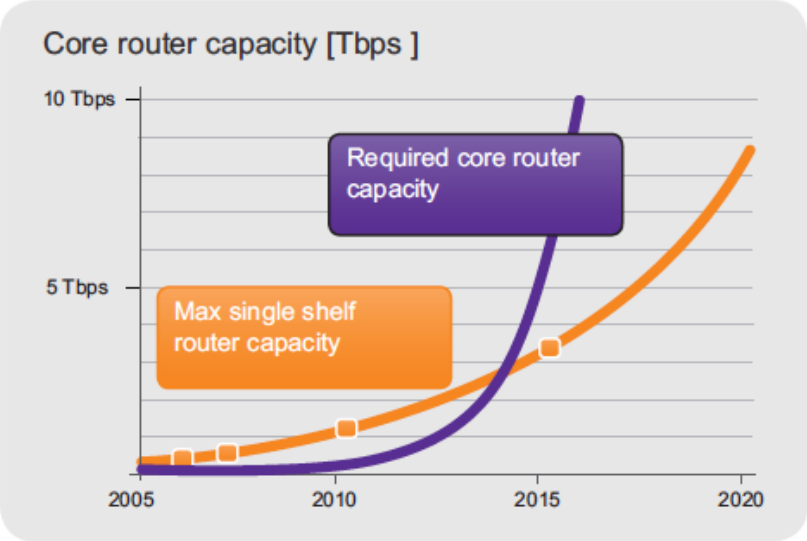


Figure 1.6: Traffic growth outpaced expansion of core routers capacity [9]

network expansion [9]. Figure 1.6 shows how the router capacity could not cope with bandwidth demand, especially as the demand is expected to grow exponentially while the router capacity has lagged in recent years.

In conclusion, the inefficiency in current core networks has been caused by the rigidity of provisioning and high cost of switching in IP routers. These inefficiencies are preventing the network from meeting the current bandwidth demands and are consuming a lot of resources. In an attempt to meet the increased demand, we are exploring techniques to improve network transportation efficiency by focusing on enhancing the router bypass process in order to reduce the general network costs and improve adaptability and agility.

1.5 Dissertation Organization

The remainder of the dissertation is organized as follows:

In Chapter 2, we explore multiple core network technologies, explaining how traditional networks work and show the strength of the optical layer in transporting large amounts of data efficiently. We concentrate on the transport layer (L2) and IP layer (L3) technologies such as MPLS, WDM, (OTN). Then, we explain Software-defined Networks (SDN) and Network Function Virtual-

ization (NFV). The automation and orchestration mainly in (SDN) will be integrated in our proposal for router bypass.

In Chapter 3, we explain the concept and workings of the router bypass, illustrating how it is provisioned traditionally. Then, we describe the main advantages of bypassing routers and show the side effect caused by fixed portioning links in a traditional router bypass. Reducing statistical multiplexing was one of the reasons why router bypass has not been widely deployed in core networks.

In Chapter 4, the impact of bypass granularities is studied in router bypass performance. We have developed ways to enhance router bypass by using finer bypassing channels; using the HAO feature in OTN, we propose techniques to make a router bypass adaptive to traffic behaviour. Unlike a traditional bypass, feedback-based logic can feed the provisioning system to enhance the bypass mechanism.

In Chapter 5, Software-Defined Network (SDN) is used as an automated provisioning system. With the assumption of an integrated network (i.e., layers 1, 2, and 3), an improvement in router bypass is being illustrated. The traffic driven network allows for more predictable provisioning which reduces the side effects of the router bypass, mainly the inefficient utilization of the links. SDN allows for a more tailored router bypass based on applications needs. We show how we can control with SDN the trade-off between maximizing bypassed traffic and link underutilization.

Chapter 6 concludes the dissertation.

1.6 Bibliographic Notes

The work in Chapter 4 was published in [66] [68], and the work in Chapter 5 is in the process of publication [71].

Chapter 2

Overview of Core Networks Technologies

In Chapter 1, we saw the main problems limiting the Internet bandwidth expansion. In order to understand the proposed solution, we need to explore the networking technologies used by SPs, which will briefly explain some of these technologies.

2.1 Network Switching Principles

A telecommunication network involves a number of terminal nodes connected links or by intermediate nodes. These links allow telecommunications between the terminal nodes. The communication can be categorized, based on various considerations; for instance, in the way of accessing the network: scheduled or random access. Another consideration is the structure of the network: tree, bus, mesh, star or ring network. In this section, we are interested in how traffic is switched from node to node in a network. The transmission links that connect these devices usually use one of two switching mechanisms: circuit switching or packet switching.

Circuit switching was originally developed to manage telephone calls over the public switchboard. In circuit switching, a dedicated connection is established between two end-systems: a dedicated non-shareable channel is reserved between the source and the destination for the duration of the connection. The

dedicated connection is established using a process called "call setup". Call setup finds the path from the source to the destination and establishes the connection between them. It can only be used for a single call at a time. The data transfer takes place only after the call is set up; after the transfer is completed, the call is cleared and the reserved resources are released [24].

IP packet-switched networks move data in separate small blocks (packets) based on the destination address attached to each packet. When they reach the destination, the packets are assembled to form the sent message. Packet switching is traditionally done by intermediate nodes called routers, transmitting the message by sending small packets along the path to be shared among all users. Every router uses a look-up table (or routing table) for each incoming packet. If a routing table does not contain the closest match to the packet destination address, the packet will be sent to the default route. The default route is either statically programmed or dynamically learned by the routing protocols. After choosing the outgoing route, the outgoing port is identified to the destination direction by looking up and routing decisions, which is computationally expensive [26].

Circuit switching is usually associated with old technologies such as the creation of the analogue telephone; nevertheless, the efficiency of transmitting large files over a dedicated path has strong potential. In contrast, the forwarding process is relatively slow when compared to circuit switching, including repeated excessive processing at each intermediate router. In addition, there is the repeated expensive process (looking up routes) at every intermediate router. Considering that the current transmitted web pages and file sizes over the Internet have been increasing in the last decade [27], switching large files in the transport layer can be more efficient alternative by using methods such as optical burst switching (OBS) or optical flow switching (OFS).

2.2 Internet architecture models

The architecture of the Internet is complex. To develop every component of the system without conflict or lack of compatibility, two network models have

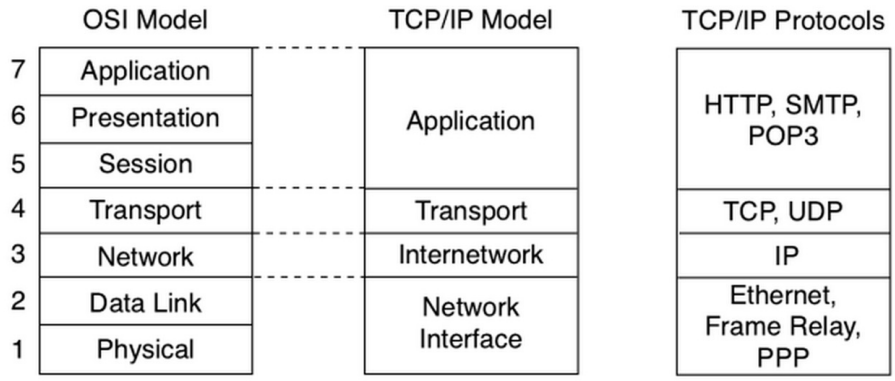


Figure 2.1: OSI and TCP/IP models [29].

been established: the OSI model and TCP/IP model.

In 1977, the International Organization for Standardization (ISO) adopted the Open Standards Interconnection (OSI), a model that breaks down many tasks involved in moving data from one computer to another. In other words, the goal of the OSI model is to break down the task of data communication into simple steps. These steps are called layers and the OSI model is made up of seven layers, each with specific responsibilities [28]. The model groups similar communication functions into one of seven logical layers. Each layer serves the layer above it and is served by the layer below it.

The Internet protocol suite (TCP/IP) is a suite of protocols used to communicate over the Internet and contains four layers. The TCP/IP model was created after the OSI seven-layer model for two main reasons: first, the foundation of the Internet was built using the TCP/IP suite and through the spread of the World Wide Web (WWW) and Internet, TCP/IP has been preferred, and second, a project researched by the US Department of Defence (DoD) consisted of creating the TCP/IP protocols. The DoD's goal was to introduce international standards that could not be met by the OSI model. Figure 2.1 shows the layers of both models.

Before briefly describing every layer in the OSI model, the cross-reference between the two models needs to be addressed. Layers 1 and 2 in the OSI model correspond approximately to the Host-to-Network layer in the TCP/IP

model. Layer 3 in the OSI model corresponds directly to the Internet layer in the TCP/IP model. Both have the transport layer on same level; the application layer exists in both models.

Since the OSI model has seven layers and the TCP/IP has four, many protocols separated in the OSI model into different layers are considered in the same layer in the TCP/IP model. For instance, TLS/SSL and FTP belong to the same layer as the TCP/IP model (application layer) while they are separated in two layers: the session and application layers, respectively, in the OSI model.

Here is a brief description of each layer in the OSI model:

Layer 1. Physical Layer: The physical layer has the following major functions: (a) it defines the electrical and physical specifications of the data connection; (b) it defines the relationship between a device and a physical transmission medium (e.g., a copper or fibre optical cable). This includes the layout of pins, voltages, line impedance, cable specifications, signal timing, hubs, repeaters, network adapters, host bus adapters (HBA, used in storage area networks) and more; (c) it defines the protocol to establish and terminate a connection between two directly connected nodes over a communications medium; (d) it also defines the modulation or conversion between the representation of digital data in user equipment and the corresponding signals transmitted over the physical communications channel. This channel can involve physical cabling (such as copper and optical fibre) or a wireless radio link. For instance, the physical layer of Parallel SCSI operates in this layer, as do the physical layers of local-area networks LANs such as token ring, Ethernet, and IEEE 802.11 and the physical aspect of modern personal communication such as Bluetooth.

Layer 2. Data Link Layer: The data link layer provides a reliable link between two directly connected nodes by detecting and potentially correcting errors that may occur in the physical layer. Point-to-Point Protocol (PPP) is an example of a data link layer in the TCP/IP protocol stack. The ITU-T G.hn standard, which provides high-speed local area networking over existing

wires (power lines, phone lines and coaxial cables), includes a complete data link layer, which provides both error correction and flow control by means of a selective repeat Sliding Window Protocol (SWP). Another example, Ethernet standard IEEE 802.3 includes this layer within its specification based on the OSI model.

Layer 3. Network Layer: This layer provides switching and routing technologies, creating logical paths (known as virtual circuits) for transmitting data from node to node. Routing and forwarding are the main functions of this layer as well as addressing, Internetworking, error handling, congestion control and packet sequencing.

Routing—part of this layer—is the process of selecting the best paths in a network. Packets are transferred from source to destination through this path. These packets may transverse cross-points until they reach the destination [30]. Based on the routing table (book of addresses and the corresponding local outgoing port) calculated by the routing protocol—an algorithm that calculates and decides the best path from source to destination—the router switches packets from the incoming port to the outgoing port.

The process that is responsible for building and calculating the routing table inside the router is called the control plane; the switching process is called the data plane. In addition to routing messages, the network may (or may not) implement message delivery by splitting the message into several fragments, delivering each fragment by a separate route and then reassembling the fragments, reporting delivery errors, etc. Datagram delivery at the network layer is not guaranteed. Distinguishing between the data plane and the control plane is important throughout this proposal.

Layer 4. The Transport Layer: The transport layer is the fundamental layer at which any end node (computer) can communicate and establish a connection with another node. The main purpose of the transport layer is to establish, maintain and release connections for the hosts involved in the communication [31]. Well known TCP and UDP protocols belong to this layer.

The segmentation of large messages to a sites lower layer (Layer 3) is one of the responsibilities of this layer, as well as reassembling the segmented packets into the original message.

Layer 5. The Session Layer: This layer governs the dialogue during communication. This layer might set up different sessions or connections. Its responsibilities are: how to establish the connections, how to use them, and how to break them down, as well as checking for transmission errors. The layer can add different headers during the data transmission [31].

Layer 6. The Presentation Layer: The main purpose of this layer is to ensure that the transmitted data being transferred is in under-stable syntax by the computer at the other end. Therefore, this layer converts the data to under-stable syntax and may encrypt and compress the data before sending it down to the session layer [31]. An example of this layer is converting an EBCDIC-coded text file to an ASCII-coded file.

Layer 7. The Application Layer: This layer is the users access to the network. The primary job of this layer is to manage communication between different applications. Some examples of this layer include FTP, HTTP, and SMTP protocols, which interact with the software running on the end computer [31].

This brief overview of OSI and TCP/IP models has been provided to show that in this thesis we will explore certain areas of the OSI and TCP/IP models, specifically, the bottom three layers in the OSI model.

Networking Technologies Overview

Most traffic in the carrier network is IP based. In addition, the evolution of IP/MPLS technology (as a flexible carrier solution implemented last decade) in the telecom network made packet switching the main way of transferring data over the telecom networks. As we described in the routing definitions, these packets need to be processed by many routers to reach their destination. In this section, we will explore many technologies currently used by carriers to introduce the concept of router bypassing.

2.3 IP/MPLS

In Multi-Protocol Label Switching (MPLS), packets are transferred based on labels. The labels have a local significance to the router itself, not to the network. The labels create a virtual circuit called a Label Switching Path (LSP).

The main idea behind MPLS is label switching. Each label is basically an integer number inserted between Layer 3 and Layer 2; because of that some researchers consider an MPLS as a 2.5-layer technology. MPLS switching allows the router to do the switching without a traditional IP lookup or searching in the routing table. Instead, the router uses forwarding information base (FIB) and Label Information Base (LIB) tables to forward the packets with a new label, which will be explained later.

Any received packets will be attached to a label, then the receiving router will recognize which output port this packet needs to be switched to. The idea is like switching in Frame Relay networks using DLCI stacks [31].

The MPLS network is built from two types of router: the edge LSR (Label Switching Router) and core LSR. The edge LSR is located as an entry point for the service provider network. The entry edge LSR is called ingress edge LSR and egress edge LSR to the exit LSR. The ingress router inserts a label (PUSH function) and the exit LSR removes the label (POP function).The routers in the middle of the MPLS network are just switching labels (SWAP

function) as shown in Figure 2.2.

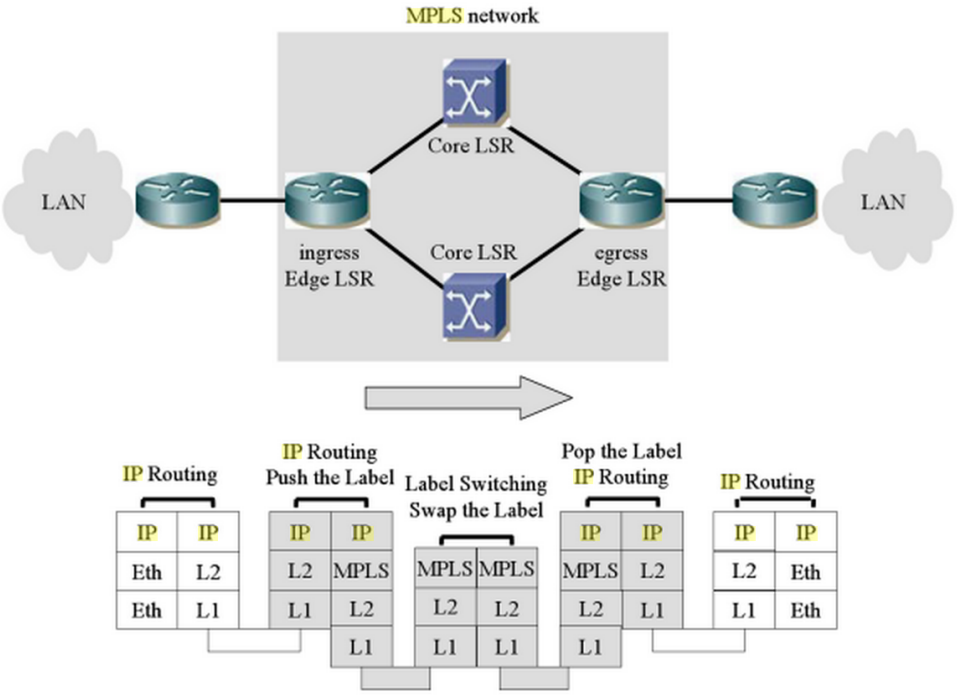


Figure 2.2: Data Transfer in MPLS [31].

One entry in the MPLS router is called Forwarding Equivalent Class (FEC). All traffic is classified into FECs. All packets sharing the same FEC will follow the same path and are linked based on the best route from the routing table. There is a unique LSP for each FEC and every LSP is one-directional so FEC classification helps to put traffic with a common LSP into one category. FECs will be a useful concept in Chapter 3, which describes the router bypassing proposal.

There are many tables inside every router where they exchange data to switch packets to the right direction until they reach the destination. After the calculation is done by the routing protocols, the best path is stored in the Routing Information Base (RIB) or the routing table. This table belongs to the control plane and it gets updated if any link changes happen to the network. Then, RIB produces the updated light version of the routing table, which includes destination addresses and the associated outgoing port, called the Forward

Information Base (FIB). The FIB is usually located in the data plane, which accelerates the switching process. In MPLS, the Label Information Base (LIB) table is populated—as well as the FIB—by the Label Distribution Protocol (LDP) or ReSerVation Protocol-Traffic Engineering (RSVP-TE). This is also located in the control plane. LIB links labels with the FEC indicated by downstream LSRs or upstream LSRs. By exchanging data between the FIB and LIB tables, a third table called Label FIB (LFIB) is built. This main table is used by the core LSRs to perform label switching (label swapping) [31].

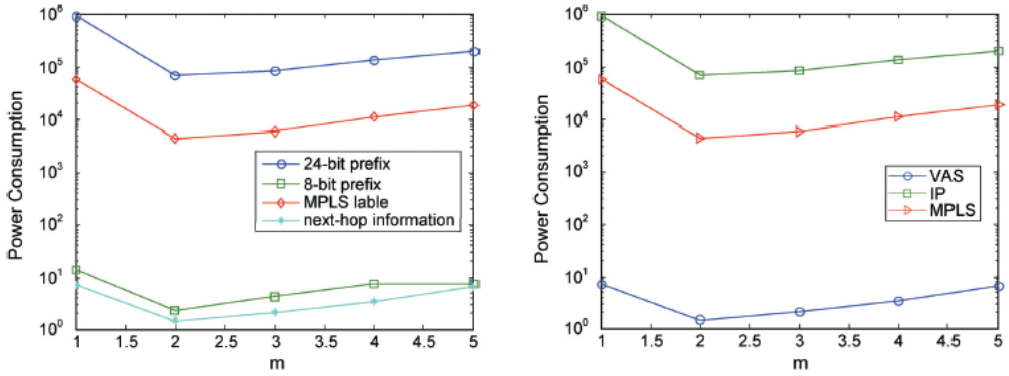


Figure 2.3: (Right)Power consumption of switching nodes(m is an integral factor). (Left) Power consumption of memories (m is an integral factor). [32].

As we have seen, the IP/MPLS is a more efficient technology in terms of switching. There is no need for IP-lookup for every packet switching. However, MPLS still consumes a considerable amount of power. The reasons are:

- the swapping process needs to be done for every packet, which includes removing and adding labels at every intermediate LSR in the carrier network. This is because the technology is based on packet switching (the data delivered based on information attached in the packet itself); and
- the size of the packets are relatively small to average.

Therefore, based on this study [32], the power consumption of switching in MPLS is less than traditional IP switching but not by a large margin. Figure 2.3 shows the power consumed by switching and memories of both technologies in that study.

2.4 Wavelength Division Multiplexing WDM

Wavelength-division multiplexing (WDM) is the process of multiplexing different wavelengths onto a single fibre. This process creates virtual channels inside a single fibre, each of them capable of carrying a different signal. Figure 2.4 shows a schematic of a bi-directional WDM system. This system has n a service interface and n a wavelength in either direction in a single fibre [33].

In WDM, data is transmitted over wavelengths either in parallel-by-bit or serial-by- character, by assigning incoming optical signals to specific frequencies (wavelengths) in a designated frequency band and then multiplexing all the frequencies into one fibre.

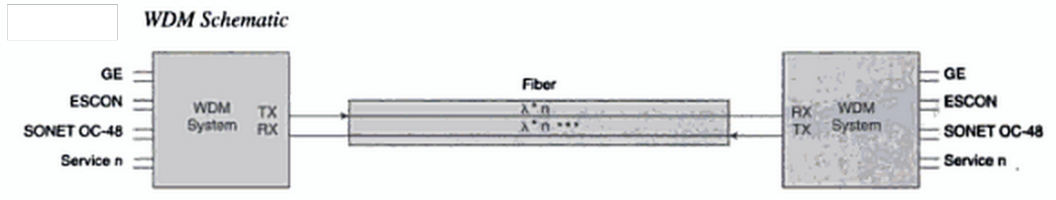


Figure 2.4: WDM schematic [33].

Each signal can be carried at different rates (OC-3/STM-1.OC-48/STM- 16, etc.) and even in a different format (SONET/SDH, ATM, etc.). WDM allows an increase in the capacity of existing networks in a scalable manner. WDM supports point-to-point, ring, and mesh topologies.

Existing fibre in a SONET/SDH fibre plant can be easily migrated to WDM. Most WDM systems are compatible with SONET/SDH short-reach optical interfaces. Long-haul WDM topologies are typically point-to-point. One of the biggest attractions of WDM is the quick deployment of new bandwidth services, which are much easier to add to WDM than installing a new physical fibre.

There are four kinds of WDM systems :

- Metro WDM (—<200 km—)
- Long-haul or regional WDM (—200 km to 800 km—)

- Extended long-haul WDM (—800 km to 2000 km—)
- Ultra-long-haul (—>2000 km—)

In long-haul WDM systems, the user service interfaces are often OC-48/STM-16 interfaces. Other interfaces are supported, including 40 Gigabit Ethernet, and 100 Gigabit Ethernet.

The transponders are the key to expanding the WDM network since they are responsible for terminating signals at a specific wavelength. Various WDM components are also integrated with WDM systems. As shown in Figure 2.5, a WDM node consists of a multiplexer/demultiplexer section, switching section and local interface section. The local interface consists of transponders and complex electronic circuitry. The transponders contain the optical source and optical detectors. The multiplexer/demultiplexer consists of an optical multiplexer and demultiplexer. The switching section usually has an Opto-Electro-Opto (O-E-O) or Optical to Optical (O-O-O) switches in add/drop configuration or cross-connect configuration. A typical WDM node can add or drop or pass through wavelengths [33].

CWDM and DWDM

There are two types of WDM: Coarse and Dense Wavelength Division Multiplexing (CWDM and DWDM).

CWDM uses a wide spectrum and accommodates up to eight channels. This wide spacing of channels allows cheaper optics to form limited-capacity channels to deliver signals over relatively short distances [34].

DWDM systems pack up to 16 or more channels into a narrow spectrum window near the 1550 nm local attenuation minimum. Since spacing between channels is smaller than CWDM, which allows the addition of more channels. DWDM requires more precise optics, which tend to be more expensive.

Typical DWDM systems provide up to 44 channels of capacity, with some new systems offering up to 160 channels. DWDM is typically used where high

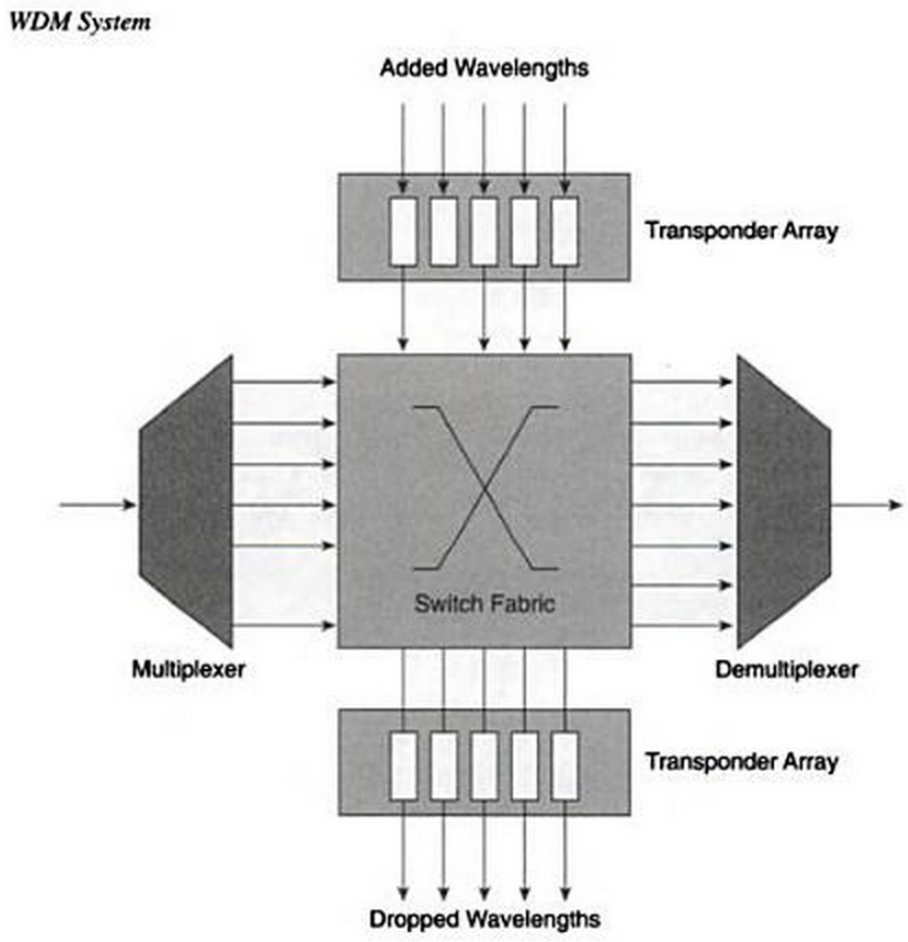


Figure 2.5: WDM system [33].

capacity is needed over a limited fibre resource or where it is cost-prohibitive to deploy more fibre [34].

ROADM and FOADM

As with most transport systems, there are different ways to add and drop traffic along ring and tapered networks. WDM systems support two types of add/drop: Fixed and Reconfigurable Optical Add/Drop Multiplexers (FOADM and ROADM).

FOADMs are based on a piece of fibre that allows add/drop of specific wavelengths. This system can be integrated and managed when cost is the important factor since they are cheaper than ROADM. **ROADMs** add the ability to switch traffic remotely from a WDM system at the wavelength layer.

Although more expensive than FOADMs, ROADMs are used in applications where traffic patterns are not fully known or change frequently [34].

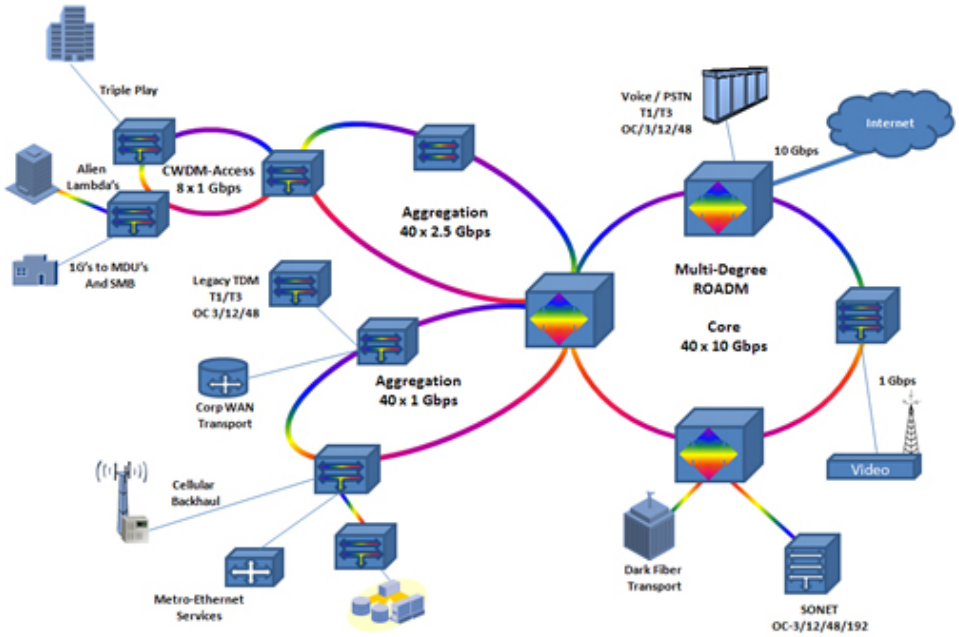


Figure 2.6: WDM network [33].

The key features and benefits of WDM include:

- *Protocol and Bit Rate Agnostic:* Wavelengths can accept virtually any services;
- *Fiber Capacity Expansion:* WDM adds up to 160X bandwidth to a single fibre;
- *Cap/Long Haul and Lo Cap/Short Haul Applications*; and
- *Remotely Provisionable:* ROADMs provide the flexibility to change with changing network requirements.

2.4.1 IP over WDM Architecture

The IP protocol is the dominant convergence and routing protocol in the current communication and networking architecture. Therefore, transporting IP traffic over a WDM network in an effective and efficient way is an ongoing

development task. Although there are many commercial products that transport IP over WDM, there are still many areas for improvement. Moreover, we are going to see the challenges facing the current IP over WDM architecture in general.

There are three known IP over WDM networking architectures, which will be explained briefly in this section.

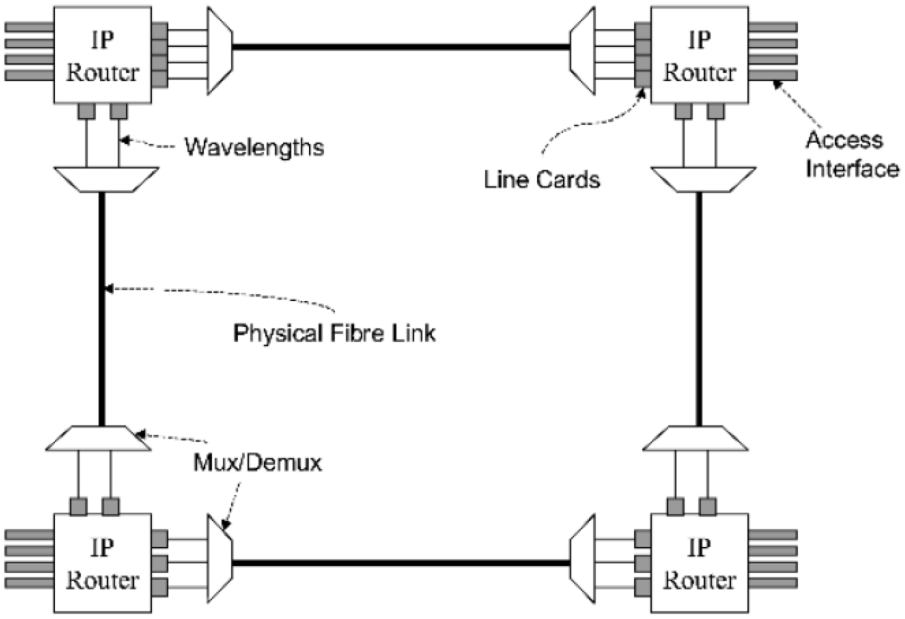


Figure 2.7: IP over point-to-point WDM [35].

IP over point-to-point WDM

In IP over point-to-point WDM architecture, IP routers are connected directly to each other via multi-wavelength fibre links. Figure 2.7 illustrates such architecture, where the neighbouring routers are connected by fixed fibre links.

WDM components such as Reconfigurable Optical Add-Drop Multiplexer (ROADM) do not form a network system but offer a physical connectivity between the IP routers. For instance, SONET can be used for framing services on the WDM channels. Packet-over-SONET can be used to encapsulate IP frames to be transported over WDM links. Point-to-point WDM systems have seen

widespread deployment in long-distance networks.

With IP over point-to-point WDM, the network topology is not easily scalable since the connections are fixed and all the network configurations are static. A centralized system typically manages such networks, with the least number of interactions between the IP and WDM layers.

IP over reconfigurable WDM

Unlike point-to-point IP over WDM, in IP over reconfigurable WDM architecture router interfaces are connected to the client interfaces of the WDM network. Figure 2.8 illustrates IP over reconfigurable WDM network as a separate physical entity. In this architecture, the WDM physical topology consists of cross-connects and add/drop interfaces with multi-wavelength fibre links. Therefore, the WDM network itself has a physical topology and a *lightpath* topology. When light crosses optical devices, it is often called the optical path.

The WDM physical topology consists of network elements interconnected by fibres, while the WDM lightpath topology is formed by wavelength channel connections. The switching in a reconfigurable WDM is a circuit-switching technology so the wavelength channel setup and termination has to happen between two ends of communication. One point needs to be emphasized here, the IP traffic switching and the wavelength switching are done in separate layers in IP-reconfigurable WDM.

Lightpaths in the WDM network are designed to comply with the IP topology. By appropriately configuring the WDM cross-connects, a given router interface can be connected to any other router interface [35].

IP over switched WDM

In an IP over switched WDM architecture, the WDM infrastructure directly supports a per-packet switching capability, as opposed to simply providing ingress-to-egress lightpaths. As such, it enables a much finer grain sharing than reconfigurable WDM. Various switched WDM approaches have been proposed, including [35]:

- Optical Burst Switching (OBS)
- Optical Label Switching (OLS)
- Optical Packet Routing (OPR)

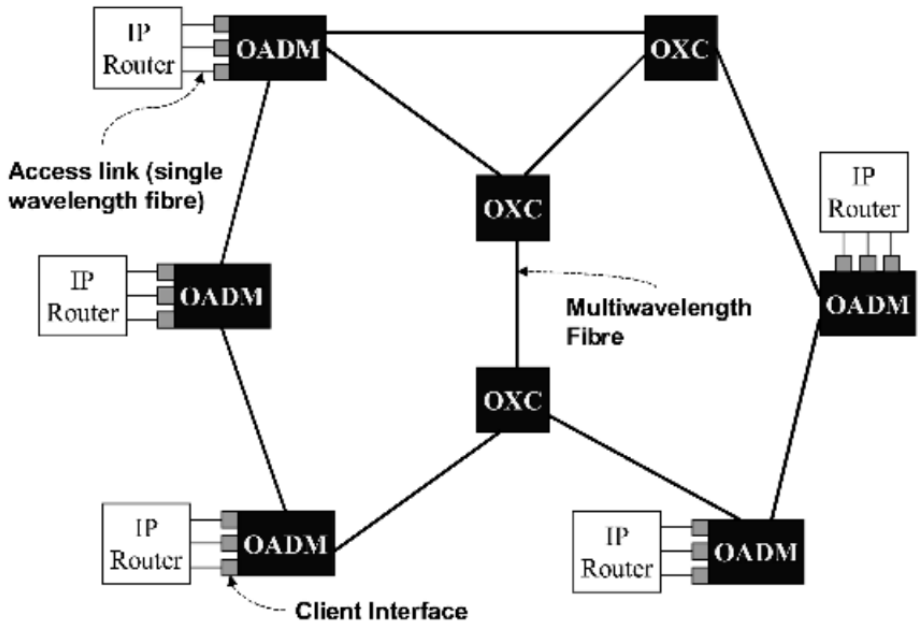


Figure 2.8: IP over reconfigurable WDM [35].

OBS and OLS offer flow-switching architecture that is like packet switching. Conventional IP packet uses packet switching based on the destination address. Switching is done in the intermediate nodes based on information in the IP header. MPLS makes the switching faster and more efficient; however, OBS and OLS (i.e., the core switches) do not recognize or use the IP header for a switching mechanism. In addition, OBS and OLS are usually meant to be used for coarse granularity flows, unlike IP packets which are very fine.

The OPR is a system where the traditional IP router is being imitated by an optical version of it. In fact, interface savings is one of the main drivers behind choosing OPR over an electrical IP router. OPR has several interfaces (i.e., more interfaces than a conventional IP router). However, because of the immaturity of optical logic processing and buffering, many of these systems cannot be fully implemented without any electronics. For instance, optical

buffers are usually represented simply by long delay lines, which are very basic and static compared with electronic memories. Even WDM systems are usually bufferless to avoid building any optical buffers inside them. In addition to optical buffers, building an optical switch fabric is a challenge in terms of having high-speed fabric while maintaining a high quality of signals. Figure 2.9 shows IP over switched WDM system.

Similarly, switched WDM systems rely on IP routers for packet switching.

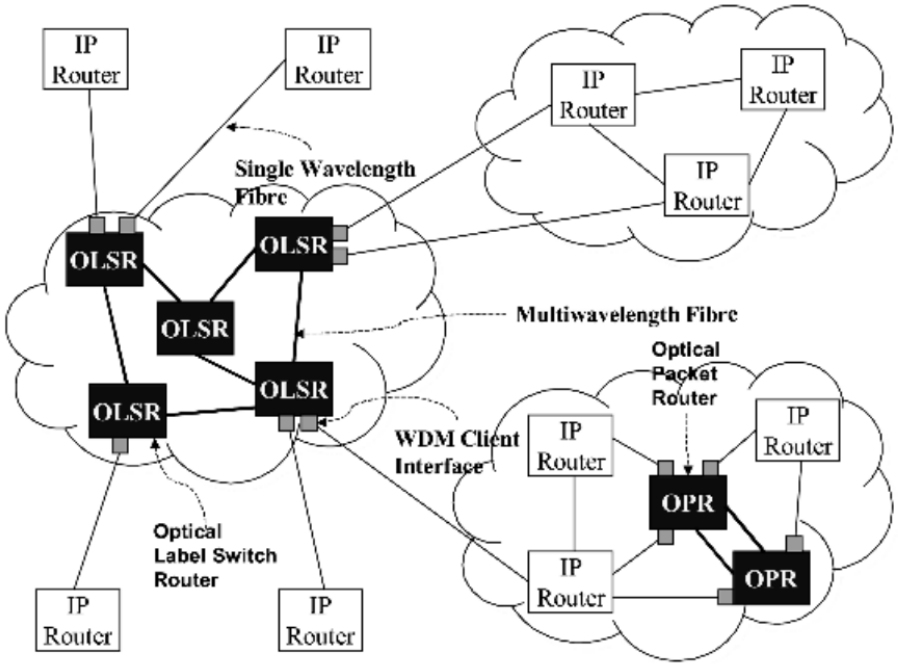


Figure 2.9: IP over switched WDM [35].

Therefore, OPR systems are not yet a finished solution compared with OBS and OLS due to implementation challenges.

The main difference between OBS and OLS is that OBS uses coarse packet switching but OLS uses application flow switching. OLS often uses an in-band sub-carrier wavelength to carry the control information, i.e., the flow header. As indicated in the figure, OLSR is formed in clusters. The Edge OLSR usually offers an electro buffering for IP packets; the edge OLSR only needs to be completely compatible with the IP protocol stack. OLSRs are interconnected

by fibres supporting multiple wavelength channels.

The three architectures presented above are associated with different hardware and control and management software. The IP over point-to-point WDM architecture will be gradually replaced by the IP over reconfigurable WDM architecture because the second architecture can offer much more functionality than the point-to-point one. The second architecture is more flexible in deployment and is scalable as well. Through carefully designed network control software and traffic engineering, the reconfigurable WDM architecture can provide much higher network resource utilization and lower operational cost than the first architecture [35].

Dealing only with lambda granularity is the main criticism of IP over WDM network. The network will be insufficient to deliver traffic because it is difficult to fill lambda with incremental increasing in traffic volume. On the other hand, the Optical Transport Network (OTN) offers sub-lambda switching (finer granularity) which is one of the main reasons for OTN popularity in recent years.

2.5 Optical transport Network OTN (G.709)

ITU-IT defines OTN as "composed of a set of optical network elements connected by optical fibre links, able to provide functionality of transport, multiplexing, routing, management, supervision and survivability of optical channels carrying client signals, according to the requirements given in Recommendation G.872" [36].

The OTN standard was developed by the ITU and is considered the successor of Synchronous Digital Hierarchy (SDH). OTN is designed with future bandwidth and protocol demands in mind, while maintaining the advantages of SDH-like flexibility and resiliency. OTN allows sub-lambda switching and offers high wavelength utilization through its flexible TDM hierarchy. It has replaced SDH as the new transport standard, enabling future proof multi-wavelength transport and management capabilities. OTN was initially used

only on point-to-point links because of its strong FEC feature but is now used as an entirely new transport network layer [37]. It is used to build transparent, scalable and cost-effective networks where existing standards such as Ethernet and SDH become client signals.

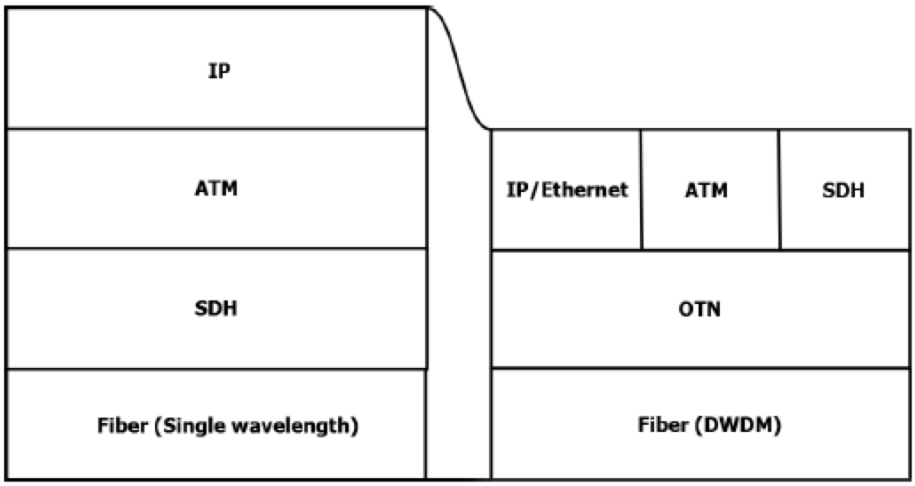


Figure 2.10: Evolution of OTN, it supports many protocols to be carried over DWDM network

Figure 2.10 illustrates the two layers: ATM and SDH collapsing into one layer: OTN as replacement. OTN is a new core transport layer and replaces SDH. SDH only supports client transport over a single wavelength and is not suitable for the management of wavelength services and addressing typical impairments in multi-wavelength optical systems. OTN accepts a wide variety of client signals (e.g., Ethernet, ATM and SDH) and provides transport and management directly over DWDM networks.

Another difference from SDH is the fact that OTN is asynchronous. This is achieved by transporting network synchronization within the payload in the OTN frame, mainly by SDH tributaries. An OTN network element thus does not require synchronization interfaces or complex clocks, helping to reduce both cost and complexity when designing the network. A drawback for operators that are upgrading their infrastructure from PDH and/or SDH to OTN is that new hardware and management systems must be installed.

Reducing the number of layers in the protocol stack has many advantages.

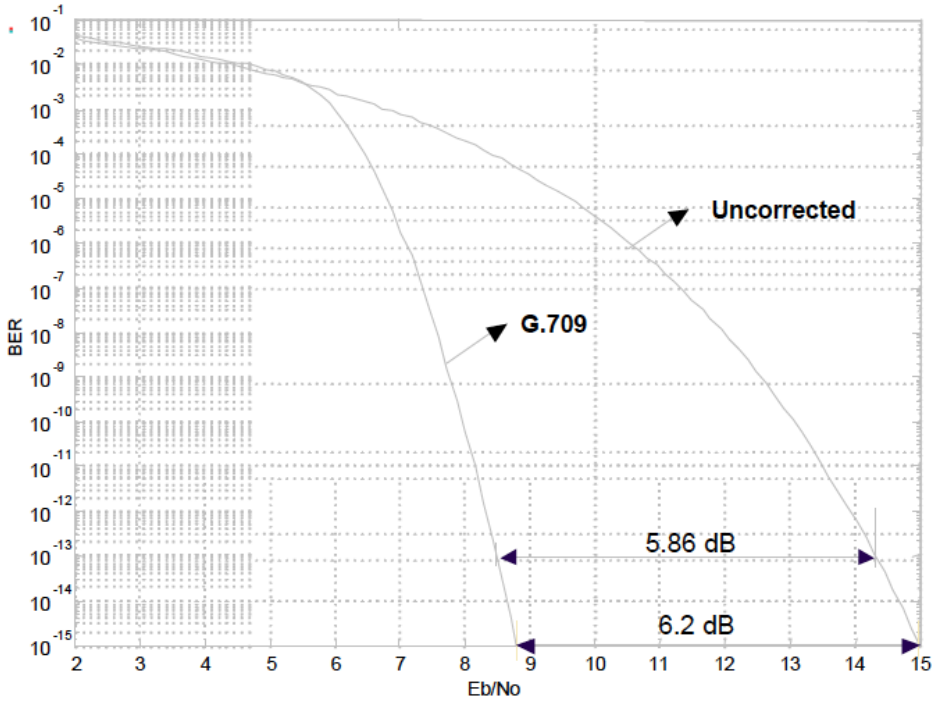


Figure 2.11: The advantage of OTN over uncorrected signal [38]

Each layer requires separate equipment, maintenance and management systems. Unnecessary duplication of functionality in the network is avoided, which saves duplicated resources. If the number of layers is reduced, the overall cost and complexity of the network will also be reduced.

One of the main strengths of OTN is that OTN uses TDM multiplexing beside DWDM multiplexing. That leads to maximum utilization of the core network bandwidth. OTN uses the capacity of each wavelength by multiplexing lower rate signals into higher rate signals. By using the TDM hierarchy in OTN, it is possible to perform sub-lambda granularity switching. It goes as low as 1.25 Gb/s in the ODU0 case. ODUflex is a flexible option for delivering to the client at variable rates. This might be an attractive and future-proof way of performing switching in optical networks. ODUflex will be explained in more detail since it is going to be part of the proposed router bypassing solution.

OTN has been introduced by the ITU as an answer for maximum utilization, manageability and wide client signal support for future DWDM net-

works. Many standardization documents belong in the OTN category. The two most important are the G.872 and G.709. The G.872 standard is called "Architecture of optical transport networks," which describes the network architecture and transport technology for the OTN, which also describes the Optical Transport Hierarchy OTH. The OTH consists of Optical Multiplex Section Overhead (OMS OH), Optical Transmission Section overhead (OTS OH) and Optical Channel Non-associated Overhead (OCh OH).

G.709 is called "Interfaces of the OTN". It defines the standard interfaces and rates for high-bandwidth optical signals, and focuses on structures, interfaces and mappings. Frame format, supported client signals, multiplexing structure, and supported signal rates are found in the G.709 standard [39].

OTN provides a flexible multiplexing hierarchy, transparent transport of client signals with backward compatibility for already used protocols, forward error correction to expand the fibre span lengths and link monitoring as shown in Figure 2.11. Sub-lambda switching in OTN is carried out with the concept of ODUs offered switching scalability which is made possible by the TDM multiplexing hierarchy. OTN switches can sometimes be a stand-alone chassis in the core switches case or even a module attached to IP routers as different media to transport traffic.

As mentioned above, OTN provides maximum use of fibre capacity by combining TDM and DWDM. TDM uses the capacity of a single wavelength by multiplexing low rate streams into higher rate streams, while DWDM uses the wide frequency spectrum in the fibre.

2.5.1 OTN frame structure

OTN incorporates a flexible multiplexing and mapping hierarchy to support a wide variety of client signals and bit rates. The multiplexing structure works like containers, accepting a wide variety of data payloads at different bit rates. Multiplexing low bit rate client signals into high bit rate signals allows for the creation of bigger data containers that are transported over a wavelength. For instance, Figure 2.12 shows how the client signal is encapsulated into ODUk

containers as well as how it gets multiplexed with other ODUk containers since every layer adds its own header.

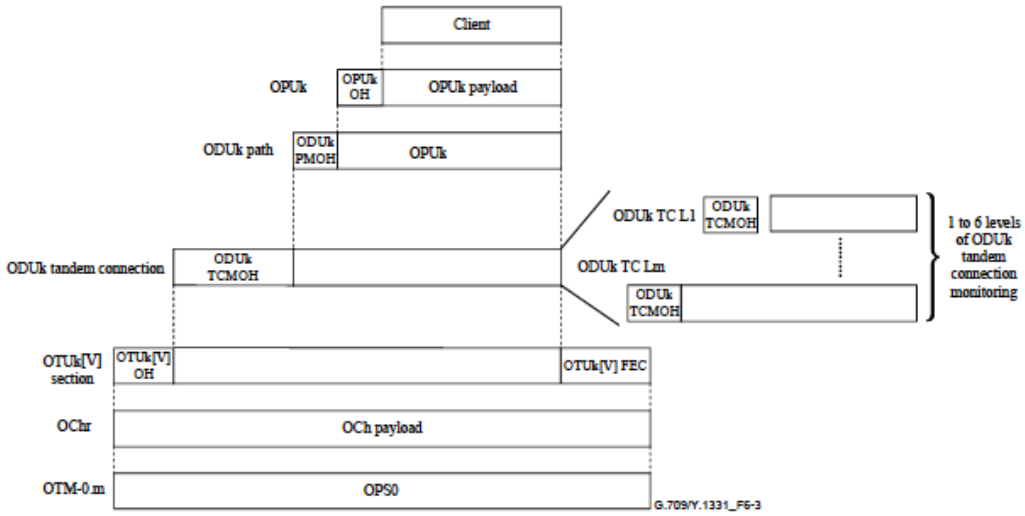


Figure 2.12: Illustration for OTN client signal encapsulation and multiplexing [39].

Figure 2.13 shows a summary of all the headers that will be added to the client signal.
Different layers have been defined by G.709 in the OTN framework as shown in Figure 2.14. The Optical Channel Payload Unit (OPUk), Optical Channel Data Unit (ODUk) and Optical Channel Transport Unit (OTUk) are in the electrical domain while the Optical Channel Non-associated OCh is in the optical domain. The details of other optical layers such as OMS and OTS are beyond the scope of this research.

2.5.2 OTN Hierarchy

The OPUk is comparable to the path layer in the SONET/SDH. OPUk encapsulates the client signal and allocates the needed rates for that signal. The ODUk function is like Line Overhead in SONET/SDH. The OTUk consists of Forward Error Correction (FEC) and performs a job like Section Overhead in the SONET/SDH. Like any encapsulation, it is added at the source and

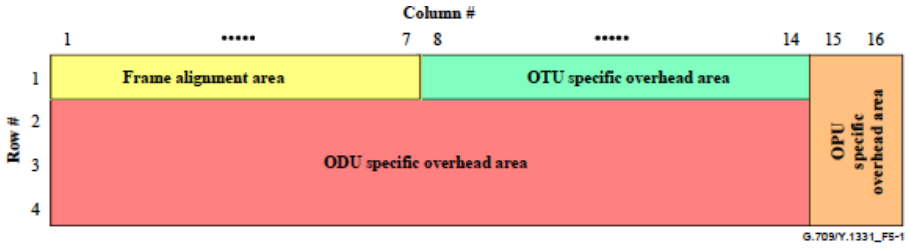


Figure 2.13: Summary of OTN overheads OPU, ODU and OTU overheads [39].

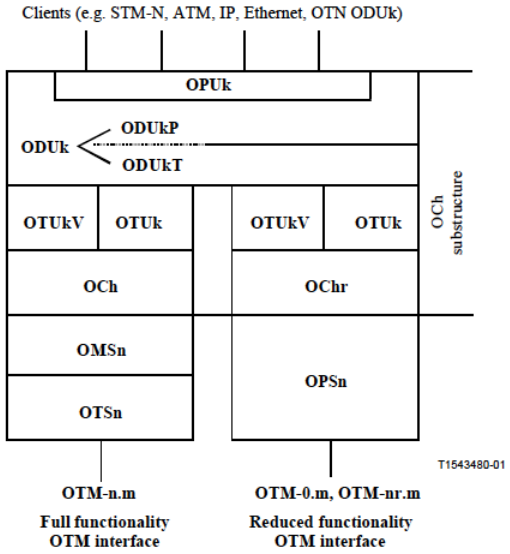


Figure 2.14: OTN hierarchy [39].

removed at the destination. After all the encapsulations have been added, including FEC, the signal will be sent into SERDES (Serializer/ Deserializer) before getting converted into the optical domain. The electrical layers will be explored further in the coming sections.

2.5.3 OTN Multiplexing

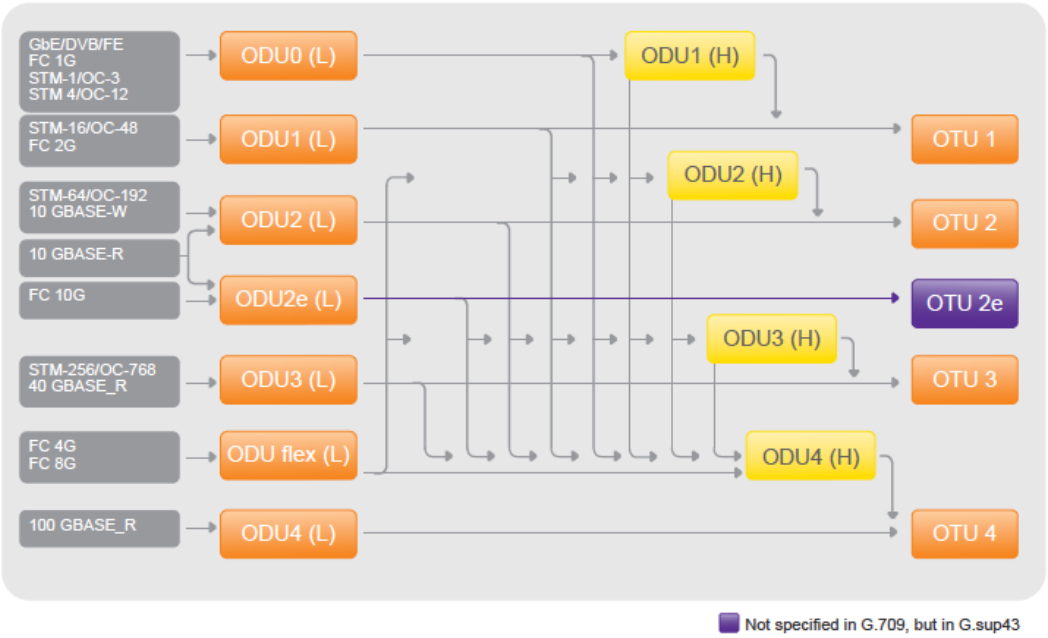


Figure 2.15: Flexible mapping and multiplexing scheme in OTN [9].

The multiplexing structure in OTN is based on putting containers into containers. Before putting the signal into a wavelength, OTN puts multiplexed data into higher rate containers. An ODU can either be put directly into an OTU, or multiplexed with other ODUs to fit inside an OTU. Figure 2.15 illustrates how client signals are multiplexed and/or mapped into different ODUs. For instance, ODU0 can be multiplexed into ODU1, ODU2, ODU3 and ODU4 where ODU2 can be multiplexed into ODU3 and ODU4. ODU2e has been designed to carry a 10 Gbps signal directly. Each one of the ODUs (except ODU0) can be encapsulated directly in the corresponding OTU.

As shown in Figure 2.15 as a more detailed view of the multiplexing, the OTUA non-OTN client signal is first mapped into a lower order OPU, named OPU(L). OPU(L) is mapped into the ODU(L), and further into the OTU[V]. OTN signals are first mapped into the ODTU in various multiples depending on the bit rate (the ODTU is an ODU with justification overhead). The ODTUs are then multiplexed into an ODTUG, which is then mapped into a higher order OPU, named OPU(H). As seen in Figure 2.15, it is possible either

to map a client signal directly or an ODTUG consisting of interleaved lower order ODUs into the OPU payload area.

The ODUk (k = 0,1,2,3,4) is used as the basic multiplexing unit in OTN. Figure 2.15 shows several multiplexing possibilities in the OTN hierarchy. It is possible, for instance, to multiplex two ODU0s into an OPU1, four ODU1s into an OPU2, four ODU2s into an OPU3 or eighty ODU0 into an OPU4.

OPU-K

The OPU accepts incoming client signals of various types. These signals can be categorized into:

- Constant Bit Rate (CBR) clients (such as ATM cell stream, SONET or SDH signals); and
- Packet-based clients (such as IP and MPLS packets, and Ethernet frames).

The OPU is responsible for mapping the client signals and adding the first layer of overhead. The OPU header consists of the Payload Structure Identifier (PSI), which includes the payload type and overhead bits associated with the mapping of client signals into the payload.

ODU-K

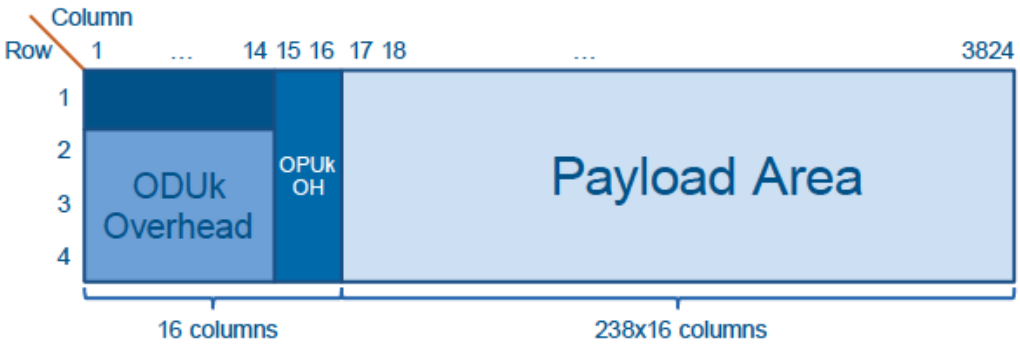


Figure 2.16: ODU frame structure [40].

Figure 2.16 shows the diagram of overall ODU frame structure.Overhead is added to the OPU, forming the ODU. The ODU overhead includes fields for operational and maintenance functions to support optical channels. Table 2.1 shows ODU types and capacity. The ODU header contains different

fields such as Path Performance Monitoring (PM), Fault Type and Fault Location (FTFL), Generic Communications Channel (GCC), Automatic Protection Switching and Protection Communications Channel (APS/GCC), Tandem Connection Monitoring (TCM) and a set of reserved bytes for experimental purposes. The ODU has several bit mapping schemes to accommodate

ODU type	ODU nominal bit rate	ODU bit rate tolerance
ODU1	239/238 × 2 488 320 kbit/s	±20 ppm
ODU2	239/237 × 9 953 280 kbit/s	
ODU3	239/236 × 39 813 120 kbit/s	
NOTE – The nominal ODUk rates are approximately: 2 498 775.126 kbit/s (ODU1), 10 037 273.924 kbit/s (ODU2) and 40 319 218.983 kbit/s (ODU3).		

Table 2.1: ODU types and capacity [39].

new clients in the ODU. The available bit mapping schemes include AMP, BMP, and GFP. AMP or BMP are used to map CBR clients into the ODU. GFP-F is used for packet-/frame-based clients. All the details of BMP, AMP and GFP are in G.709 standard [39].

An important feature in the ODU overhead is the TCM fields, allowing up to six different monitoring levels across operator domains.

OTU-K

The OTU consists of an ODU with additional overhead. The OTU is the lowest layer belonging to the electrical domain. It prepares the data to be ready to be transported over the optical layer. Fields found in the OTU header are: Section Monitoring (SM), General Communication Channel 0 (GCC0), and two bytes that are reserved for future international standardization. FEC is added at the end of the OTU frame.

FEC is a very important feature for long-haul optical transport networks. It allows for longer signal spans without signal regeneration. It is a critical feature especially where submarine fibre is being used since it is installed over long distances and it reduces the need for signal regeneration. Active regeneration is a difficult and costly process for a transatlantic communication

cable where FEC can help to reduce the need for the regeneration process (see Figure 2.11).

OTU signal rates

OTU type	OTU nominal bit rate	OTU bit rate tolerance
OTU1	255/238 × 2 488 320 kbit/s	±20 ppm
OTU2	255/237 × 9 953 280 kbit/s	
OTU3	255/236 × 39 813 120 kbit/s	
NOTE – The nominal OTUk rates are approximately: 2 666 057.143 kbit/s (OTU1), 10 709 225.316 kbit/s (OTU2) and 43 018 413.559 kbit/s (OTU3).		

Table 2.2: OTU types and capacity [39].

Table 2.2 shows the line card rates of standard OTU. OTU0 (1.25 Gb/s) is not in the table because it is not defined as a line card speed; instead, it is used as a multiplexing entity (ODU-0).

2.5.4 ODU-flex

It is hard to predict future client signal bit rates. In addition to Ethernet, many other client signals must be supported by OTN (such as Fibre Channel and video distribution signals). Most of these would not fit into any existing ODUk without significant loss of bandwidth. Defining a new ODU container for each new client signal that should be supported was considered impracticable [40].

To ensure flexible and easy adaptation to support future client signal rates, an extension known as ODUflex has been introduced to the OTN standard. ODUflex is a flexible lower order container that can be right-sized to fit any client rate. It does this by occupying a minimum number of time slots in the higher order ODUk for accommodation of the client signal. To clarify how ODUflex works, we need to mention that ODUflex is used in two cases:

- *ODUflex (CBR) for Fixed Rate Client Transport*: Most non-OTN client signals that are transported in the OTN are Constant Bit Rate (CBR) signals. The signal can be synchronous or asynchronous. In asynchronous cases, the rate is independent from the ODU rate. In this

case, a mapping agent is needed to fit and adapt the client signal into the ODU. In the synchronous case, the nominal rate of the client is same as the ODU. Therefore, no extra mapping is needed [41] [43].

- *ODUflex (GFP) for Packet Flow Transport*: ODUflex is used to transport packet-based signals. The rate of these packets fluctuates over time; therefore, a mapping mechanism is required. The Generic Framing Procedure (GFP) is typically used. The packets need to be encapsulated by GFP framing and then mapped into an ODU. Another important point is that the client flow is not encapsulated directly inside the ODUflex container like ODUflex (CBR). Instead, ODUflex (GFP) rates are multiples of approximately (ODU0) 1.25Gbit/s to correspond to the capacity of an integer number of a higher order ODU Tributary Slots, and the packet flow is adapted to that rate using GFP [43]. Therefore, the size of ODUflex must be N x ODU0 where N is an integer number.

2.5.5 Hitless Adjustment ODU-flex

In late 2011, the HAO standard was published as a draft for G.7044 with a period of study of around four years. Hitless Adjustment of ODUflex (GFP) (HAO) is a resizing mechanism in the OTN that allows it to support an increase or decrease in the ODUflex (GFP) client data rate across its entire end-to-end path. HAO provides a control mechanism hitlessly to increase or decrease the capacity of an ODUflex (GFP) connection to meet the bandwidth needs of the application [44].

To achieve a hitless bandwidth adjustment of an ODUflex (GFP) connection, all switches need to support the HAO standard otherwise the connection must be torn down and re-established. The bit rate adjustment of the ODUflex (GFP) occurs simultaneously among all the nodes in the ODUflex (GFP) connection to prevent buffer overflow or underflow [44].

As mentioned above, it is possible to change the rate of an ODUflex (GFP) in order to adapt to changes in traffic patterns. Although there is no explicit limit to how frequently an ODUflex (GFP) is resized, the motivation for devel-

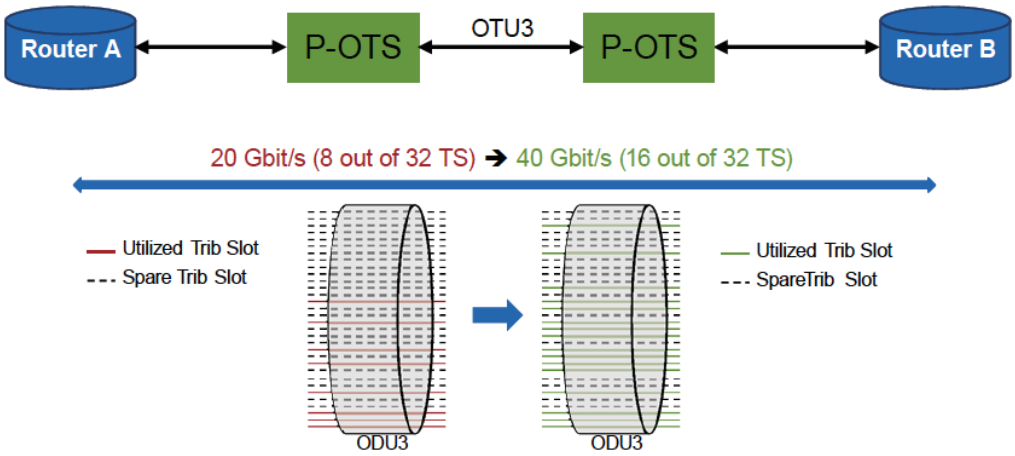


Figure 2.17: ODUflex (GFP) resizing example [43].

oping a resizing protocol was to address longer-term changes that occur over a period of months or years.

To change the size of the ODUflex end-to-end connection, the number of tributary slots should be changed from link to link and even in the switch connection with the switch for the whole connection to be adjusted. In the case of an increase in bandwidth, new Tributary Slots must be allocated prior to the change in rate of the ODUflex (GFP). In the case of a rate decrease, the ODUflex (GFP) must decrease in rate before the Tributary Slots are freed up.

Consequently, there are two aspects to the HAO protocol. One is the Link Connection Resize (LCR) and the other is the Bandwidth Resize (BWR). Link Connection Resizing is triggered at each node by either the management plane or the control plane. Both node-to-node and end-to-end handshaking concepts exist, ensuring that the number of Tributary Slots for the entire connection has been properly adjusted and that the correct Tributary Slots are being used. Bandwidth Resizing is controlled by the source node in each direction. The source node either increases or decreases the rate of the ODUflex (GFP). One field (called GMP) changes at the source node which leads all intermediate switches to follow that adjustment.

Figure 2.17 shows a simplified example of an increase in bandwidth between

routers supported by a change in rate of the ODUflex that carries the packet data. The source router requests an increase in the bandwidth to double (40 Gbps), which is OTU3. Then, the two nodes agree; therefore, another eight tributary slots are allocated to the connection and the ODUflex increases the bandwidth to meet the expansion request.

2.6 Comparison

Restoration feature	IP/MPLS	OTN	WSON
Degree of automation	Automated	NMS or with GMPLS automated	NMS or with GMPLS automated
Layer	Any granularity	Sub lambda services	Wavelength
Degree of protection	Can protect only high priority traffic	Protects per sub lambda service	Protects whole link
Need for termination?	Yes – on router. In IP over DWDM all packets are processed by the router	Yes – on ODU XC	No
Time to restore	Potentially minutes Using IP, < second for MPLS-based techniques	< second	Longer than OTN < IP MPLS

Figure 2.18: Comparison of restoration techniques in networks [5].

Figure 2.18 shows the link restoration difference between the networks mentioned above. IP/MPLS can process very fine granularity traffic. However, that comes at a cost in processing in terms of power, as well as the IP network taking the longest period for restoration compared with other networks. On the other hand, the IP over WDM network offers very coarse switching granularity.

OTN offers a medium ground between the two technologies (WDM and MPLS) in terms of granularity with a much faster restoration mechanism. The flexibility and speed of restoration (assuming the provisioning is the same) has been one of the motivations for this proposal.

OTN will be used in our solutions in Chapter 4 and 5 mainly for the following reasons:

- Wide range of bandwidth granularities which allows us to establish multiple sizes of bypass channels based on the need.
- Fast link restoration.
- Ability to change the bypass channel size on the fly without tearing down the connection.
- Optical bandwidth will be groomed and managed effectively.

2.7 Software Defined Networking (SDN) and Network Function Virtualization (NFV)

Software Defined Networking (SDN) and Network Function Virtualization (NFV) are recently introduced technologies designed to tackle many insufficiencies in network architectures and operations. We will explore these technologies and how they might enhance router bypassing in general.

2.7.1 Software Defined Networking (SDN)

SDN has been defined as an "the physical separation of the network control plane from the forwarding plane, and where a control plane controls several devices" [46]. Others define it as "emerging networking paradigm that may allow changes in the limitations of current network infrastructures. First, it breaks the vertical integration by separating the networks control logic (the control plane) from the underlying routers and switches that forward the traffic (the data plane). Second, with the separation of the control and data planes, network switches become simple forwarding devices and the control logic is implemented in a logically centralized controller, simplifying policy enforcement and network re-configuration and evolution" [45]

From these definitions, it can be seen that the core concept in SDN is about separating the control plane (the network brain) from the data plane

(the physical equipment). In traditional networks, the distributed control plane on every physical device led to rigid complicated systems which made it hard to adapt to the changing application environment.

2.7.2 Strength of Software Defined Networking (SDN)

SDN transforms fragmented network devices into one coherent platform which offers adaptive network services to deal with the volatile demands of the application layer. SDN facilitates faster new organization application deployments, drastically cutting the operation cost by enabling policy-based workflow automation. SDN technology enhances cloud architectures by offering automated, on-demand application delivery and mobility of scale. SDN expands the concept of data centre virtualization and enhances resource efficiency and utilization by reducing infrastructure costs and overhead.
SDN achieves these enterprise goals by forming a centralized management and orchestration platform which automates the provisioning and configuration of the whole network infrastructure. In traditional networks, IT groups (i.e., Network, Data, application, etc.) are normally separate and require extensive time to communicate to each other before deploying any changes to their IT infrastructure. In SDN, common IT policies bring IT groups together faster by simplifying network services and reducing the complication of deployment details. That led to advance network infrastructure capable of offering new application services in minutes compared to the time needed by traditional network operators.

SDN delivers speed and agility when deploying new applications and business services. SDN vendors are promising to deliver a flexible, policy-based programmable platform which is capable of handling most of the network demands in the future. The following are the main benefits of deploying SDN:

- **Enable innovation**: providing open network services will pave the way for innovations for new businesses and applications.
- **Offer new services**: helping telecom companies to generate income by offering new network services.

- **Reduce capex**: installing more hardware in plug-and-play fashion and configured just by zero-day setup

- **Reduce opex**: supporting automation and orchestration will drastically reduce costs associated with network operations. Traditional network require human interaction for any change in the network. In SDN, most of the day-to-day configuration will be automated within policy-based centralized controllers. Support of end-points grouping enables IT operators to segment their network privileges according to the business policy. For instance, by assigning any device into the finance department, then that device will receive the privileges and security policy of that department without human involvement.

- **Deliver agility and flexibility**: supporting business to provide new application and services will help them to be more competitive in the market place. Fast network provisioning and providing more network services will offer better visibility for the network resources, allowing them change with changing network circumstances. For instance, in the case of congestion in a certain area of the network, the application will have an abstract view and will be able to choose other routes.

- **Programmability**: the capability to use open APIs to connect applications to the network. Traditional networks lacked any common API which made it hard to program an application which interacts directly with network resources. Programmability is an essential tool to customize applications based on business needs [46].

2.7.3 Basic SDN architecture

An SDN architecture consists of three layers: the application layer, controller layer and infrastructure (or physical) layer. The application layer is where, for instance, the provisioning portal is located, network virtualization, network monitoring, intrusion detection (IDS) and flow balancing and other applications. Next, the essential part of the SDN is the controller layer. SDN controllers are formed by decoupling the intelligent part (control plane) from the physical equipment. The network global intelligence is formed on software

rather than in the confinement of the fragmented and separated hardware. Controllers are integrated with physical devices which allow them to interact with all equipment by pushing requests to the devices. At the same time, controllers also collect data from the devices to form an abstract view of the status of the network. That abstract view facilitates the application layer to use network services more intelligently and efficiently [47].

The SDN architecture is based on the following fundamentals [46]:

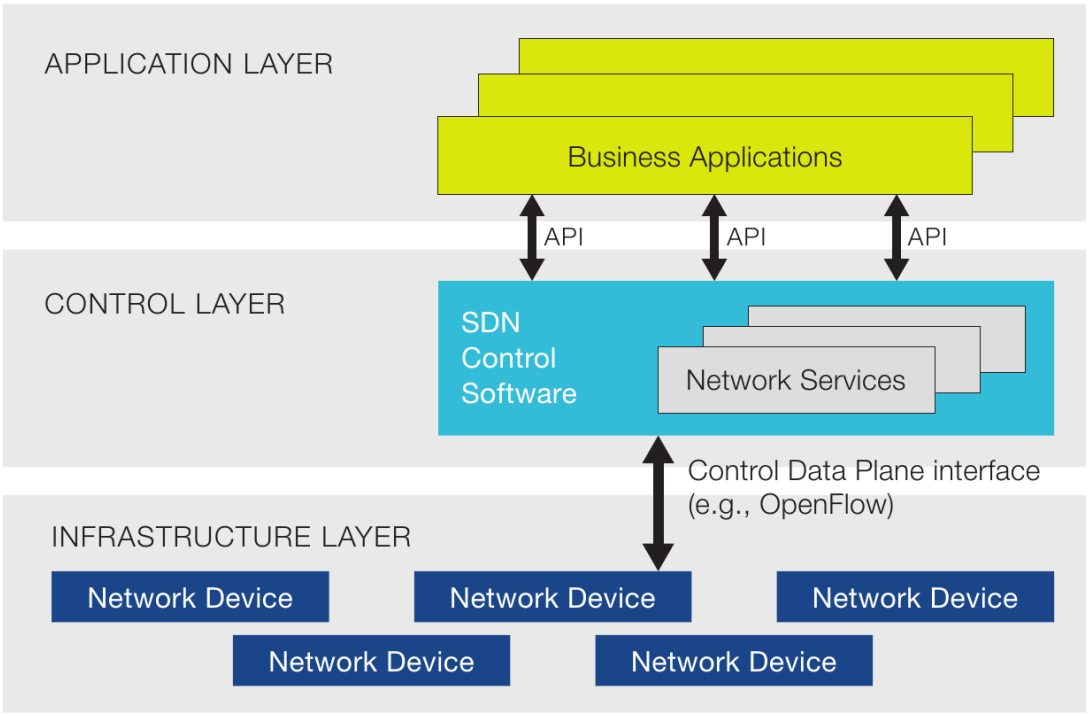


Figure 2.19: Principles of SDN architecture [48].

- **Directly programmable**: Since the network control plane is separated from the forwarding plane, it allowes the network to be directly programmable.
- **Agile**: Simplified abstraction offered by network controllers has paved the way for more agile and adaptable network.
- **Centrally managed**: centralized software-based SDN controllers maintains an abstract and global view of the network. That abstraction offers

the application a simplified network view to choose from the most appropriate and efficient network services.

- **Automated operations**: SDN allows network operators to configure, manage, secure, and optimize network resources in a fast and dynamic fashion using automated SDN policies which can be enforced on any white box (no proprietary-based devices) to carry out the changes needed [46].

  Building SDN networks based on open-standard technology allows controllers to communicate with every device in the network. Vendor neutrality simplifies network day-to-day operations and design because instructions will be the same for all devices even when manufactured by different vendors.

OpenFlow has been adopted by many well-known vendors as an essential part of SDN programming. OpenFlow was created at Stanford University and is now being managed by the Open Networking Foundation (ONF). The objective is to build a new common language to communicate with network switches. OpenFlow is used mainly as a communication tool between controllers and network devices. That communication allows controllers to collect data from the devices and at the same time force the way traffic flows should be routed in these devices.

OpenFlow became popular first with service providers, including Google, which encouraged many vendors such as Alcatel-Lucent, Brocade, Cisco, Dell, F5, HP, Juniper Networks, NEC, Plexxi, and VMware to adapt and support it as members of the ONF [47].

2.7.4 SDN using APIs

Application programming interfaces (APIs) are used to build highly programmable infrastructure which will produce the abstraction needed in the SDN framework. As a communication channel, APIs are used to send instructions to program a device. Developers code their application based on APIs documentation to write the appropriate command where the device will understand

and respond to it. The SDN framework includes southbound APIs and northbound APIs depending on where that communication tool is located.

APIs located on controllers will be used by applications to send requests. That communication channel through API is called the northbound API. APIs located on physical network devices and used by controllers. These APIs are used to send requests to network equipment to implement the changes needed. The communication API between controllers and network devices is called southbound API [47].

2.7.5 SDN network overlay

To avoid the massive cost associated with replacing traditional network equipment, network overlay is proposed by SDN as an option to carry out SDN implementation while minimizing the cost. Installing virtual switches in place of existing network switches will allow this equipment to communicate with SDN infrastructure seamlessly. By creating hypervisor (separate OS) inside the switch, the switch will able to create an API tunnel to communicate with controllers. These tunnels will be used to pass instruction on how the device should route traffic.

Many protocols have emerged including VXLAN, STT, and NVGRE to make the network overlay feasible by using network encapsulation.

2.8 Network Function Virtualization (NFV)

Network Function Virtualization (NFV) is a method to reduce the cost of network deployments by separating network functionality (i.e., firewall, switching, routing, etc.) from any specific hardware. Instead, that functionality can be virtually hosted on a server.

Cutting hardware costs is accomplished by using open-based hardware without vendor or specific hardware restrictions. Telecom operators will be able to host multiple functionalities on the same physical hardware. Therefore, Telecom companies will not only avoid the cost of expensive hardware-specific

equipment but will be able to use one physical device for multiple functions. For instance, if a customer requests switching, routing, and firewalling services then the telecom operator—with NFV—can accommodate all the services on a single physical hardware system. Hosting multiple customers on the same hardware using separate virtualized instances will also be possible with NFV [49].

Besides cutting costs, NFV adds flexibility and opens doors to more innovations. Depending on open-based hardware instead of proprietary dedicated hardware will allow network operators to move network functions around the network based on demand. That flexibility is difficult to achieve without NVF. Decoupling the network's functions from the hardware will pave the way for more innovations by experimenting with new networks functionalities without the high cost of hardware deployment. Simply by spinning a new virtual machine (VM) into existing hardware will allow a new network service to be installed and offered for use.

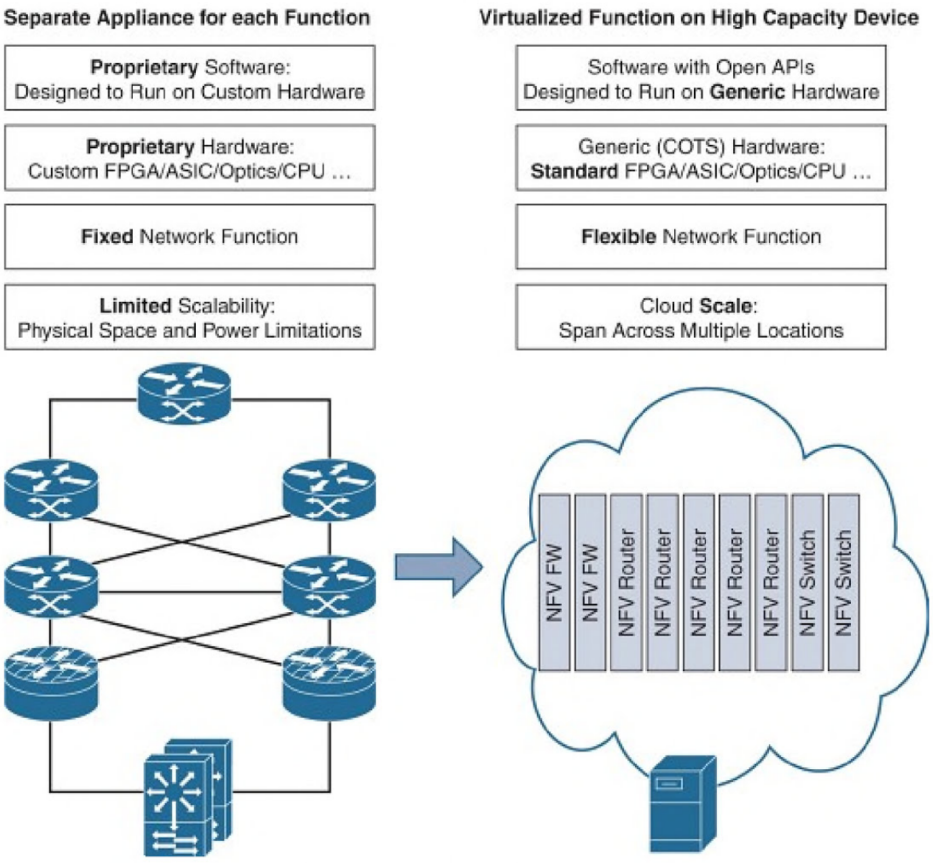


Figure 2.20: Transition into NFV [50].

2.8.1 The Advantages of NFV

NFV uses industry-standard hardware and NVF relies on software for network services instead of dedicated hardware. By using software, network deployments and management will be faster and simpler. Adding and removing network services and even troubleshooting can be done from anywhere within the network.

The concept of separation functions from hardware in NFV has many benefits, which include:

- Saving costs in hardware
- Cutting costs in hardware space
- Cutting costs in power consumption
- Cutting costs in hardware maintenance
- Easier and faster network maintenance
- Using hardware more efficiently

2.8.2 SDN and NFV

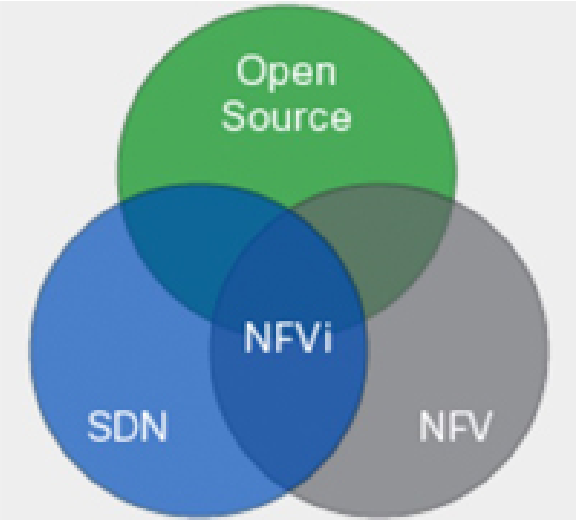


Figure 2.21: SDN and NFV [51].

Both technologies (SDN and NFV) are designed to increase flexibility, reduce costs, enhance scalability, and speed up the introduction of new services.

They are separate technologies and one can be deployed without the other. However, the technologies complement each other. SDN makes using NFV much easier; centralized orchestration allows the network to deploy services based on NFV in a more efficient, flexible and simpler fashion. SDN also helps NFV address tasks such as policy-managed forwarding and dynamic service orchestration. In turn, NFV will facilitate the operation of SDN; in particular, NFV paves the way for traffic switching (router bypass, traffic engineering, etc.), dynamic scale-up and scale-out, multi-tenancy and load balancing. For instance, if the SDN application new is ordered—even if it does not exist in the current network—NFV will allow for faster accommodation by spinning the new services in software. The use of dynamic virtual overlays and need for multi-tenancy in NFV also leads to higher demands for SDN.

2.8.3 ETSI Framework for NFV

NFV was initially proposed at the SDN OpenFlow World Congress in 2012 by main service providers. They summarized the major issues faced by network operators–mainly the issue of being trapped within the proprietary hardware made by vendors. The following are some of the challenges:

- Design changes around the new equipment
- Deployment cost and physical constraints
- Need for expertise to manage and operate the new proprietary hardware and software
- Dealing with hardware complexity in the new proprietary equipment
- The short lifecycle that makes this equipment rapidly obsolete
- Restarting the cycle before the returns from the capital expenses and investments are fully realized

The group proposed NFV as a way to tackle these challenges and improve efficiency by leveraging standard IT virtualization technology to consolidate many network equipment types onto industry—standard high—volume servers, switches, and storage. These could be in data centres, network nodes and in the end-user premises. To realize this goal and define a set of specifications that would

make it possible to move from the traditional vendor and network-centric approach to an NFV-based network, seven of these leading telecom operators formed ETSI Industry Specification Group for Network Functions Virtualization (ETSI ISG NFV) group.

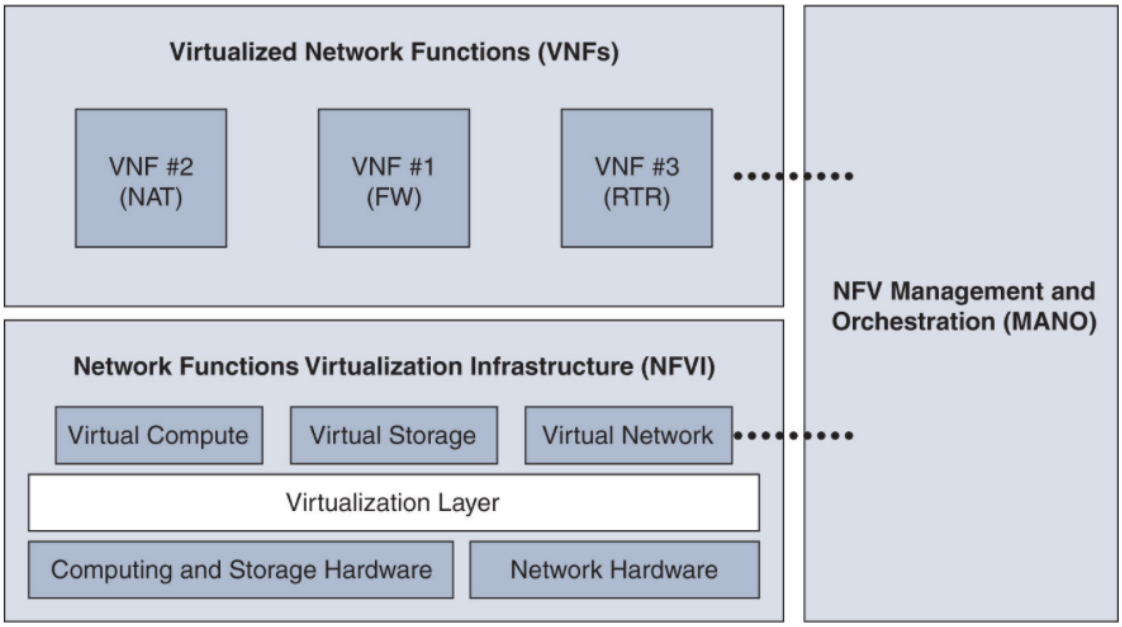


Figure 2.22: NFV framework [50].

The group used three key criteria as recommendations:

- Decoupling: complete separation of hardware and software
- Flexibility: automated and scalable deployment of the network functions
- Dynamic operations: control of the operational parameters of the network functions through granular control and monitoring of the state of network

Based on these criteria, a high-level architectural framework was established, defining distinct areas of focus as shown in Figure 2.22.

In Chapter 5, we will be exploring SDN framework in router bypass context. Provisioning network based on applications demand is a favourable way to improve router bypass. SDN automated framework will help to reduce the side-effect of traditional router bypass.

Summary

In this chapter, many transport and networking technologies have been explored as the background for this proposal, most of which will be the basis for the proposed router bypassing concept. MPLS protocols as a transport method have already been implemented by many carriers. MPLS runs over L3 routers which uses a packet switching mechanism. WDM and OTN as transport layers (L2 and L1 technologies) have basically been used to create point-to-point links between L3 routers. Introducing OTN will empower the transport layer to be more than just point-to-point links. Features such as the customizing capacities of the links in ODUflex and hitless adjustment for the links will enhance the transport layer to be flexible enough to handle the traffic behaviour changes as well.

Chapter 3

Router Bypass

As shown in Chapter 1, the amount of Internet traffic has increased dramatically in recent years. Therefore, Service Providers (SPs) are striving to reduce the cost of network operations in order to compete in the open market.

Any attempt to reduce the cost of operations will be appreciated from many perspectives: driving down the cost of bandwidth to customer and reducing the carbon footprint by reducing power consumption.

3.1 Cost of IP electronic routers vs. optical switches

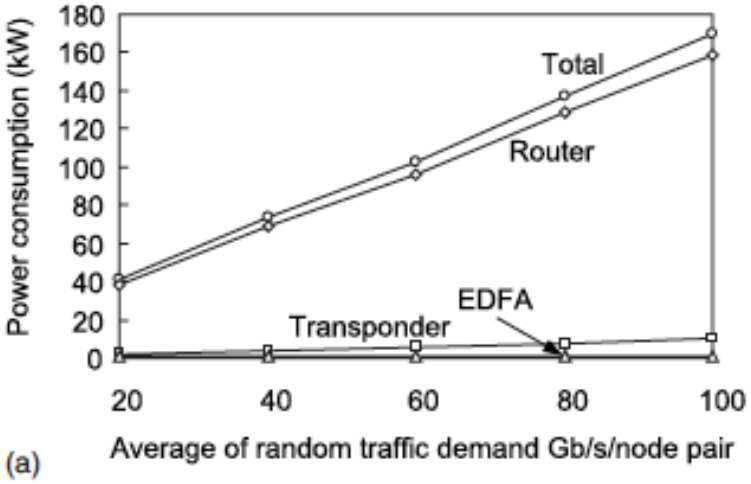


Figure 3.1: Power Consumption of IP Routers vs. Optical Switches [8].

As stated in the previous chapters, the power consumption of L3 IP routers is much higher than L0 and L1 optical switches. Many studies show that power consumption is one of the bottleneck issues in expanding the Internet [8]. For instance, an IP over WDM networks pictured in Figure 3.1 shows the huge difference in power consumption between the routers and the optical switches. The cost of current core IP routers is high compared with Optical L0 and L1 switches and they consume more power. As shown in Table 3.1, router ports are more expensive than optical ones. Nevertheless, SPs are adding more of these routers in the core networks to keep up with bandwidth demand. In the case of an IP over WDM network, IP routers consume 90% of the total network power consumption:transponders and EDFAs consume much less power to occupy about 7% and 2% of the total power, respectively [8].

Router (40 Gb/s port)	$80,000 per port
40 Gb/s transponder	$25,000
Optical EDFA	$1,500
Optical multiplexer/demultiplexer	$5,000

Table 3.1: Estimated costs of router ports vs. optical ports [8].

3.2 The concept of router bypassing

The idea of router bypassing or off-loading means the transit traffic skips being processed by transit L3 routers; instead, the traffic will be switched by lower-layer devices (i.e., L0 or L1 optical switches). Figure 3.2 explains the idea of router bypassing. Router R1 is sending traffic to R2 within the core network. Considering that the average transit traffic on core routers is about 70% [52], approximately 30% of that traffic will be terminated and processed by R2. In the bypassing design, the transit traffic (around 70%) will be directed by the optical or OTN switch to R3 directly. Without bypassing, the traffic needs to be processed by R2 and the transit traffic will be routed to R3.

To understand how bypass can be deployed in the MPLS network, the packets are labelled to be switched more efficiently as explained in Chapter 2. In MPLS, in contrast, all traffic with the same label will be assigned to a specific

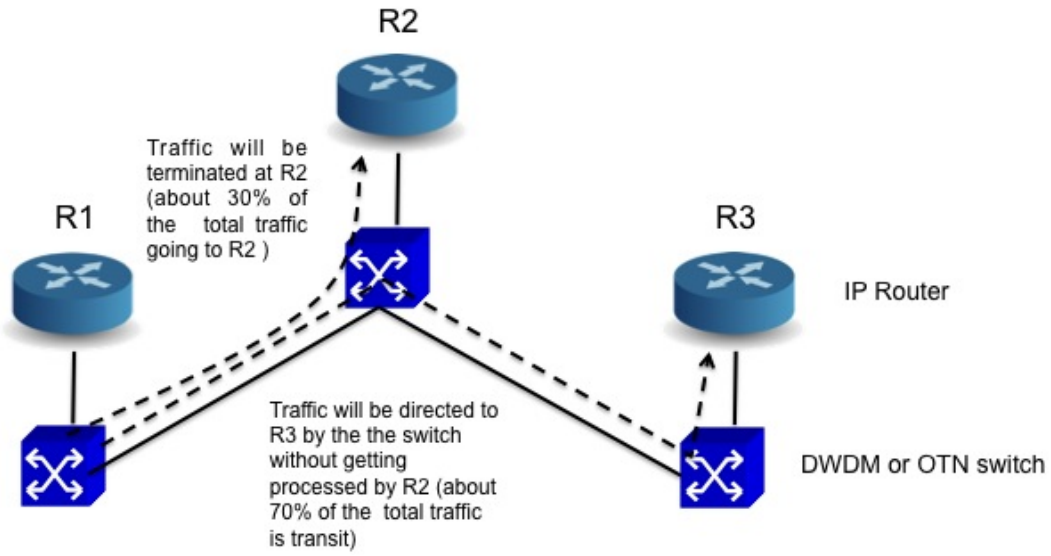


Figure 3.2: The concept of router bypassing or off-loading.

frequency (lambda). Therefore, the specific paths of any given labelled packets will be assigned to the optical path. That path will be managed by optical switches without the need for a transit router to process that traffic.

3.3 Previous Work

Many studies have explored the bypassing concept in different contexts. For instance, in the IP-over-WDM context, a study [8] showed the power savings that can be achieved by router bypassing techniques. However, because of the coarse granularity of WDM and the other reasons we have seen in the last chapter, the IP-over-WDM network is not a popular solution for router bypass; therefore, the idea of implementing router bypassing on top of WDM will not go very far. Another study [53] proposes a mathematical model for router-bypassing without considering explicitly the underlying technology which does not offer a practical approach for implementation.

A recent study [54] suggested an approach based on IP and dynamic-circuit-switching (DCS). Even though they have shown results which appear superior to router bypassing ones, the approach is impractical. First, the IP and DCS techniques require eliminating the packet switching between the core routers which requires replacing all the core and border routers. That will not only

remove the statistical multiplexing advantage of MPLS completely in the core, but any connection established in the core network requiring a bandwidth less than 1.25 G bps ODU0—the minimum size of OTN channels—will under-utilize the network. Secondly, replacing core routers with new hardware is not an incremental solution which can be deployed on top of existing technologies.

Therefore, offering solutions to improve the network performance and, at the same time, build on top of the existing technologies is important. Accordingly, our approach is built on top of OTN technology which has been growing significantly among carriers in recent years - for its resiliency and other reasons which have been shown in Chapter 2. In addition to that, we exploit the new introduced features of OTN such as HAO into router bypassing techniques in order to enhance the network performance and resolve some of the issues facing a router bypassing approach.

3.4 ways of deployment

In this section, we will describe router off-loading which can be implemented in many ways. We will focus on:

- Router off-loading using the OTN layer.

router off-loading by the OTN layer

In this technique, the bypassing is carried out by the OTN/ ODU cross-connect switches. The transit traffic will not pass to the intermediate routers, as shown in Figure 3.3. Using OTN/ODU, the cross-connect will replace the legacy SDH cross-connect; therefore, the network will be able to support higher Ethernet speeds. It will also add some of the router's functions to the same OTN physical box. OTN bypassing can be implemented in different scenarios. One scenario is that the OTN switch will be mapped into one router or two routers as shown in Figure 3.3. The router deals only with local traffic while OTN switches deal only with transit traffic and pass the local traffic to the routers. Another scenario is where the OTN network co-exists with the MPLS network. They are both connected to the edge routers. The OTN network will be

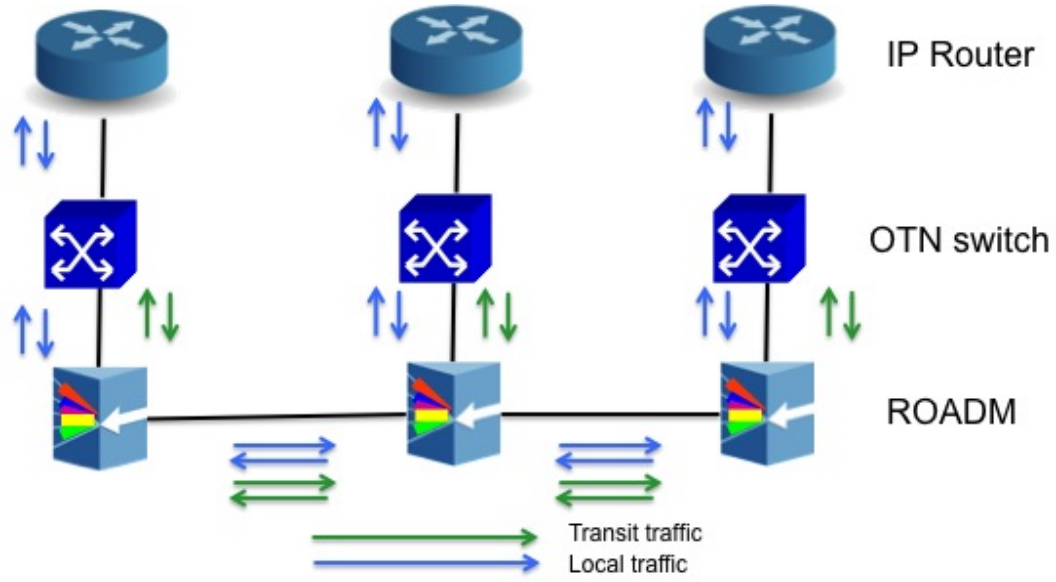


Figure 3.3: Router off-loading architecture by using OTN switches.

responsible for transit traffic and the MPLS network will be responsible for local traffic. However, this implementation involves challenges before it can be implemented in real life at least for a couple of years. [5]. The main advantage of OTN in bypassing is that the transit can use a sublambda (ODU) switching advantage on the OTN network. The wide range of ODUs rate (containers) offers considerable flexibility and finer granularity for bypassing designs. That granularity will be discussed in upcoming sections.

Router off-loading by IP/WDM layer (optical bypassing)

In this method, the transit traffic will be in the optical layer only, as shown in Figure 3.4. The optical switches will transfer the transit traffic by assigning it to a certain lambda (frequency).Then, the optical cross-connect will establish the optical path from the source to the destination. The granularity here will be very coarse. A signalling or management system can establish the path. In IP over the WDM networks, OTN switches will not be used. WDM switches are connected directly to IP routers on a one-to-one basis.

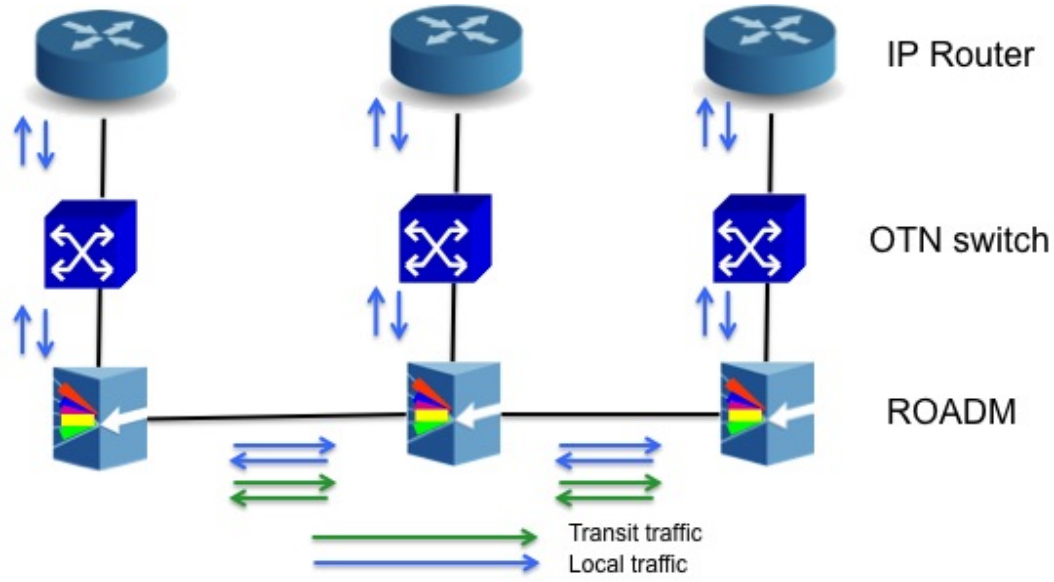


Figure 3.4: Router off-loading architecture by using WDM layer.

3.5 Router off-loading expected benefits

From a theoretical perspective, we will have the following advantages in router bypassing design:

- **Total power reduction**: As shown in Chapter 1, the cost of packet processing (IP Forwarding) is the most expensive process inside the routers. Therefore, the power needed by the transit router to process the transit traffic, which will bypass that router, will be saved. Based on study [8], in an IP over WDM network, power consumption will be reduced by 25% to 45% by using a router bypassing design.

- **Lower cost core network deployment**: Since more optical ports will be used by the bypassing mechanism for any given traffic, we will need fewer expensive ports for IP routers.

- **Better availability**: The availability of IP routers is usually around 99.9% while the optical switches are 99.999%. That means relying on optical switches (by bypassing traffic through them) will reduce down-time periods in the operator network. This 100-times factor will be significant over a long period [5].

- **Enhanced scalability**: The throughput capacity of current DWDM

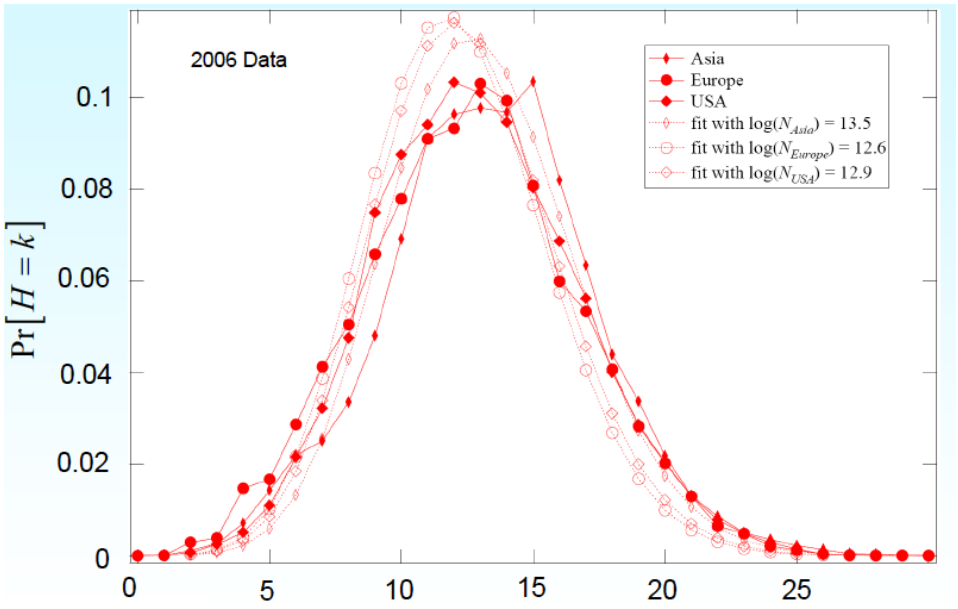


Figure 3.5: Probability distribution of hope count in the Internet [57].

and OTN switches is almost twice that of IP routers per chassis. In 2011, the capacity of OTN switches and IP routers was 8 to 9 Tbit/s and 4 to 4.5 Tbit/s, respectively [5]. In 2013, the capacity of OTN switches and IP routers was around 21 Tbit/s [55] and 12.5 Tbit/s [93], respectively. If half of the traffic is bypassed to routers, that means only half the number of IP routers are needed to handle the same amount of traffic.

- **Roadmap towards complete optical data-plane core networks in the future**: by pushing the data-plane functions toward the optical part of the network it will lead to a complete optical data plane network design and potentially a substantial reduction of the bandwidth cost.

3.5.1 Potential savings

Based on the previous bypassing scenario with some estimated value of cost per bit inside routers, this section will show the potential savings achieved by bypassing. Figure 3.5 [57] shows that most of the traffic around the world goes through 10 to 15 routers before it reaches its destination. Therefore, if we take that into consideration, bypassing a larger number of hops will save power.

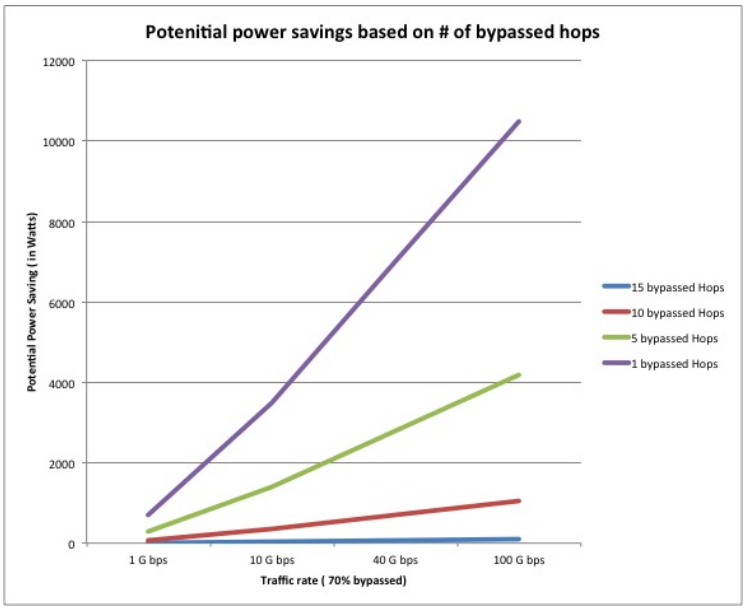


Figure 3.6: Potential power savings increases with number of bypassed hops.

Figure 3.6 shows how much power approximately (for one link) can be saved given that the energy cost of 1 bit through a core router is 10 nJ [58]. This graph shows that when we do router bypassing over high-speed links (i.e., 100 Gbps ODU4), the savings will be much higher. OTN bypassing can be done over ODU4 (about 10 Gbps), since it can carry ten ODU2 (about 10 Gbps each).

Positive Feedback: Early Router Bypassing Deployments

Router bypassing may not be the final answer for all service providers since every SP has a different situation. However, it will save a great amount of core network expansion CAPEX up to 40%, and up to 60% of power and space related OPEX [5].

Many positive testimonies from service provider have shown the positive impact of implementing router bypassing technologies in their networks. For instance, one of the first operators in Europe to adapt to the router off-loading concept, according to the company's senior network architect, noted: "The total capex for the network is 30-40% lower than it would have been when implementing conventional Layer 3 IP switching" [5].

In another example, one Asian incumbent operator interviewed by Innovation Observatory has taken the decision to move to an OTN core using a router off-load strategy. The company says that implementing OTN will mean a requirement to put in more total capacity than it would need if it were to employ routing at Layer 3 (because of the loss of statistical multiplexing), but the company is not concerned as it views it as cheaper to "waste" space in an OTN network than to expand its Layer 3 capacity to cope with traffic growth. Moreover, the company sees other important benefits for key services. According to the company's senior network architect, services such as IPTV and VoD (standard, and increasingly, HD) benefit from "a much improved viewer experience." He attributes this to the fact that IP (Layer 3) protection takes longer than 200 ms, whereas Layer 0 / Layer 1 protection switching can be achieved in less than 50 ms which he says "approaches the level of persistence of vision."

3.5.2 Drawback of Router Bypass: Reducing Statistical Multiplexing with Link Partitioning

Dedicating a certain percentage of the total bandwidth to form a coarse light path between two routers will decrease the statistical multiplexing. That might lead to two main problems: (1) less bandwidth use, and (2) greater delay for the traffic because the flow is uncertain and random. It is impossible to predict the destination of all the traffic over a period of time. Therefore, the time needed to form a light path should be justified by the amount of traffic that will use that percentage of total bandwidth. Ideally, the use should be 100%. That is one of the main criticisms of the router bypassing concept. However, these problems might appear more often on edge routers. In the core network, the router will be connected to a few routers with high-speed ports (up to 100 Gb/s) so the chance that a large amount of traffic will not be local for the next hop is high. The expectation of transit traffic in core networks is about 70% of the total traffic.

The discussion on the pros and cons of router bypassing can be considered as an extension of the debate between packet switching and circuit switching. The packet switching has a complete statistical multiplexing feature but

it is an inefficient way of transporting traffic because of the requirements for expensive and power-hungry electronic switching equipment. As a reminder, in packet switching, every packet needs to be switched individually on every intermediate node until it reaches its destination.

In contrast, circuit switching is an extremely efficient way of transporting traffic because it can use current optical technologies. Despite that fact, it still has the drawback of dedicating a certain bandwidth to specific traffic, which will degrade the bandwidth use. To minimize the statistical multiplexing issue, the bypassing channel's size should be less than the amount of traffic needed by that channel by a certain margin. How much that margin would be depends on the network operator's circumstances and the traffic patterns going through the network. These channels need to be fully used to maximize the overall bandwidth utilization. In addition, since OTN HAO allows us to change the rates on the fly, at least in theory, having a feedback function for the OTN switches will help in resizing the bypassing channels based on incoming traffic requirements.

3.6 proposing the adaptive router bypassing network

As shown in previous sections, limiting the statistical multiplexing in current packet switching networks is the main concern with router bypassing. In this proposal, we will try to introduce a mechanism whereby we can deploy the router bypassing concept with maximum utilization. In addition, we will explore ways to enhance the performance of router bypassing by supporting the network to be more traffic-aware. For instance, if it is being chosen to bypass the intermediate routers, we ask which type of traffic will enhance the overall efficiency of the core network.

The adaptive router bypassing technique can be implemented in OTN by using standard Hitless Adjustment ODU-flex HAO. The idea of HAO is that the OTN ODU rate can be changed (bandwidth decreased or increased) on the fly without the need to establish a new ODU. That will allow us to build a net-

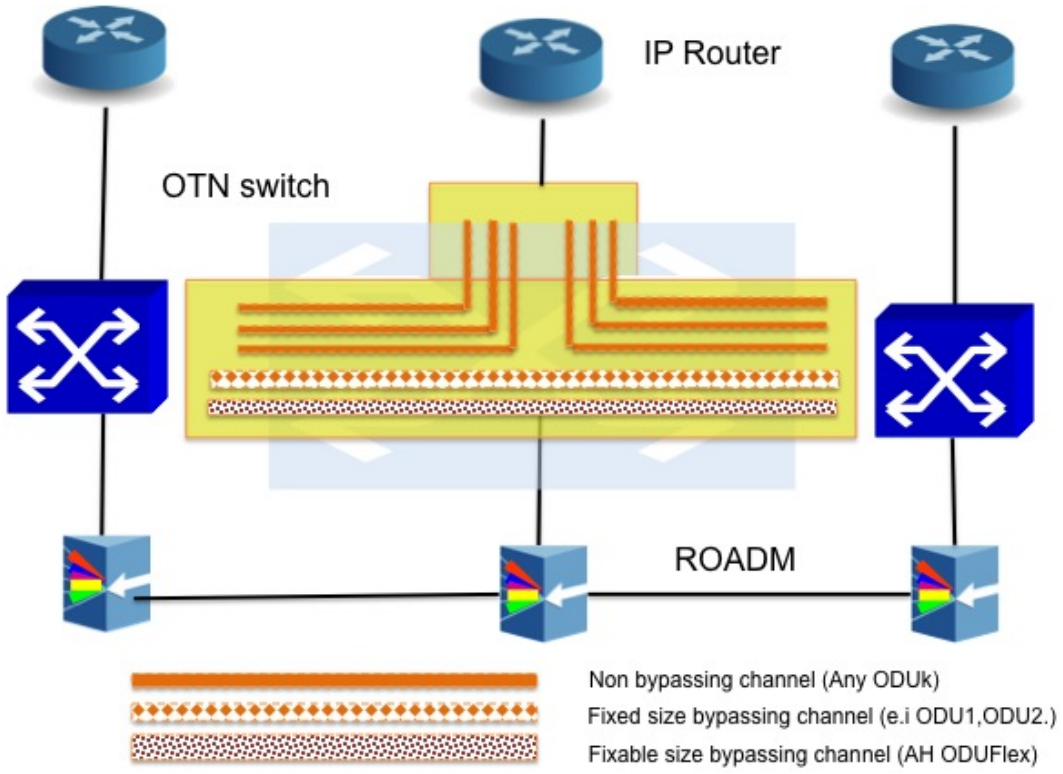


Figure 3.7: Illustration for resizable channels in adaptive router-bypassing network.

work based on the OTN and to design an adaptive router bypassing technique by using HAO containers based on the traffic pattern.

Figure 3.7 shows how an adaptive bypassing network can be implemented. Since this network is based on OTN, the OTN switch will handle (multiplex and demultiplex) the incoming data in ODUk form. By integrating the MPLS protocol with OTN switches, the OTN switch will be sufficiently intelligent to recognize MPLS paths, formerly called label switched path (LSP). By developing a certain method it can be decided which of these ODU channels will be forwarded to the intermediate router or bypass it.

Replacing the labelling mechanism in traditional MPLS, which is packet-by-packet based on channel switching is a method known as MPLambdaS switching. As mentioned above, similar traffic that would traditionally share the same LSP will be assigned the same channel (ODU in OTN).

In the adaptive routing bypassing network as shown in Figure 3.7, the bypassing channels can be fixed-size channels based on ODUk or adjustable-size channels based on ODU-flex. In addition to ODU-flex, HAO will add flexibility and allow these channels to be adjustable for their rates. Non-bypassing channels will be forwarded to a local router for conventional traffic processing. Traffic can be local or transit. In that way, the network will add the bypassing feature side-by-side with traditional router processing.

Of course, the system will be limited by implementation challenges. For instance, two main challenges involve in the adaptive design: (1) the speed of reconfiguring OTN switches: if reconfiguration is relatively fast it will help the bypassing performance, (2) traffic volatility: how much and how fast traffic behaviour changes over a period. Adapting to the real time traffic is virtually impossible, therefore, prediction or another mechanism is needed to optimize the bypassing capacities.

In this thesis, the following ideas will be explored to establish an adaptive (dynamic) router bypassing approach.

3.6.1 Granular Bypassing

One of the main issues with router bypassing is losing statistical multiplexing to a certain extent, caused by the uncertainty of the incoming traffic. Therefore, we propose to have a back-off degree of bypassing or soft bypassing. Soft bypassing is where the capacity of the bypassed links is less than the capacity of the total traffic volume. The extra traffic will be sent over the normal link (statistical multiplexed) to be routed by the transit router. The result will be less queuing time because more capacity will be available in the normal link to transmit more packets in the buffer. Also, we will guarantee that the bypassing link is fully used. The back-off value will depend on the probability distribution of incoming traffic.

$$C_{bp} = \alpha \times V_{trans} \ , \tag{3.1}$$

where C_{bp}: capacity of the bypassed link, α: percentage value and V_{trans}: volume of the transit traffic. V_{trans} will be all traffic that falls under the same FEC. Traffic sharing the same FEC uses the same LSP and that is exactly what is required.

3.6.2 Adaptive bypassing link based on traffic behaviour

The unpredictable behaviour of the Internet today has been one of the challenges for the infrastructure operator. At peak hours of Internet use, traffic increases by 25% [12] of the average traffic. Moreover, expanding the network infrastructure based on peak hours is not feasible economically for the service providers (SPs) especially when the return per bit has been declining for the last decade.

Based on this proposal, the issue might be solved by building an adaptive reconfigurable network. For instance, by using HAO, any bypassing channel can be increased or decreased, based on the traffic requirements, for a period of time. In addition, the flexibility of the bypassing channel's rates can be any rate (i.e., N x ODU0, where ODU0 is around 1.25 Gbps). Therefore, the size of the bypassing channels will be adjusted to maximize the efficiency of the transport network with the least effect on the bandwidth utilization.

The right volume of traffic to establish a bypassing link will be related to the volume of the transit traffic and the power consumption of the OTN network and routers per bit which can modelled as the following:

$$V_{trans} > \beta \times C_{bp} \times \frac{T_{otn}}{T_{Router}} \times \frac{P_{otn}}{P_{Router}} \ , \tag{3.2}$$

where β: a scaling factor (dimensionless), T_{otn}: time needed to reconfigure OTN switches, T_{Router}: time needed to reconfigure router (change in routing table to occur) P_{otn}: power cost per bit in OTN switches and P_{Router}: power cost per bit in IP routers.

We can draw from the previous model many factors that control the vol-

ume of transit traffic required to establish the bypassing link. First, the time to reconfigure OTN switches and in routers should be considered. The time of reconfiguration in OTN switches is in seconds while changing the routing table will destabilize the whole network which takes up to 10 - 100 X the time needed by OTN switches to stabilize. Second, the power that is consumed by ports in routers is is higher than that consumed by optical ports in OTN switches.

3.6.3 Content-based router bypassing

The concept of content-based router bypassing is to make the network aware of the transported traffic. Certain types of traffic for router bypassing is an area (to our knowledge) not yet examined in research papers. In content-based router bypassing, we can prioritize traffic for the router bypassing process in such a way that it will enhance the overall efficiency of the transport network. For instance, small-sized packets (i.e., less than 100 bytes) require more switching resources than larger packets for a given amount of traffic. If this kind of traffic can be bypassed more often, we will save more resources, assuming there are no side effects by doing that. This assumption will be investigated further during this research.

For bypassing to be implemented, more intelligence is required (traffic engineering features) for OTN switches to assign certain traffic to a specific channel size. Fortunately, a lot of traffic engineering work has already been done in MPLS. Therefore, using MPLS traffic engineering features over OTN will eliminate any extra work needed to build a content-based router bypassing network. However, that does not mean work is not needed on MPLS to consider the OTN infrastructure in that sense.

Cross Multilayer Switching Capable Network

In traditional networks, L3 IP routers function regardless of the underlying transport network. The L3 layer has its own mechanism for bandwidth provisioning and restoration, and does not exchange a lot of information with lower layers. Building networks that give different parts of the network resiliency can

be developed without affecting other layers. For instance, the same IP router can be used whether the underlying transport network is OTN or SONET. Although the transport network has developed a new protocol, the L3 IP router has not been directly affected, which does not mean the different layers cannot share information to build a more efficient integrated network. The transport needs to have some intelligence, which the IP router has, to operate more efficiently.

IP routers build complex routing algorithms by exchanging messages. Some of this collected data is important for the transport network as well. Unlike optical switches, IP routers have powerful computing resources. Their functions are based on the operating system software OS, so adding extra function to any router can be achieved simply by updating the OS. As an example of the importance of some data for the underlying transport network, collecting data for the traffic pattern (volume of traffic toward a destination) over a period in core routers will provide a foundation for building the bypassing channels in the underlying OTN network. Researchers have recently been calling for a united integrated network [5], which is not far away in practical terms, especially with some new commercial routers which have the OTN switching functionality in the same chassis [59].

3.7 Preliminary Simulation

A simulation has been built to test the proposed thesis. OMNET++ is the main simulation environment used for this work. OMNeT++ is an extensible, modular, component-based C++ simulation library and framework, primarily for building network simulators [60] [61].

In addition, the INET framework has been used. The INET framework is an open-source communication networks simulation package for the OMNeT++ simulation environment. The INET framework contains models for several wired and wireless networking protocols including UDP, TCP, SCTP, IP, IPv6, Ethernet, PPP, 802.11, MPLS, OSPF, and many others [60] [61].

3.7.1 INET Submodule

Many different compound modules have been used from the INET suite to build the router bypassing simulation. Each of these compound modules consists of simple modules integrated to function as one module.

- **StandardHost**: IPv4 host with SCTP, TCP, UDP layers and applications. Basically, it is a UDP packet generator with built-in packets sinks for the received packets. Many parameters can be specified in the initiation file to control the speed of the generated packet, burstiness of the generated traffic, packet sizes, UDP destination port and more.

- **Router**: IPv4 router that supports wireless, Ethernet, PPP and external interfaces. By default, no wireless and external interfaces are added; the number of Ethernet and PPP ports depends on the external connections.

  The router submodule is being used as a router with a static routing table. The routing table has been built separately in *.xml file. Static routing is being used to offer full control over the routing process in the simulation. The type of routing protocol comes under the control plane tasks related to this area of research.

The simple modules from the INET suite include the following:

- **IPv4NetworkConfigurator**: This module assigns IP addresses and sets up static routing for an IPv4 network. It assigns per-interface IP addresses, strives to take subnets into account, and can also optimize the generated routing tables by merging routing entries.

  The configurator supports both manual and automatic address assignment, and their combinations. One can provide address and netmask templates with unspecified ports, and the configurator automatically completes them by trying to put nodes on the same LAN into the same subnet. It also supports manual routes, and automatic routes that follow the shortest paths. By default, the configurator adds default routes where applicable (e.g., in hosts) and does subnet-based routing [60] [61].

In this simulation, all the auto static routing has been turned off. Instead, static routing has been specified manually on every router, which gives control over the routing process instead of the default routing created by the configurator.

The used packages are the following:

- **Channel DatarateChannel**: The channel package is used to specify the parameters of the connection between the devices in the simulation. The parameters are propagation delay, data rate, bit error rate (BER) and packet error rate (PER).

3.7.2 Building a simulation for router bypassing

This simulation is built to emulate the traffic and switching mechanism of the router bypassing scenario. Packet delay, packet loss, bandwidth use and queuing time inside each router has been measured to calculate the difference between the traditional network and the bypassing case. Moreover, we will attempt to improve the router bypassing concept by considering mainly adaptive router bypassing.

Figure 3.8 shows the scenario of our simulation. There are two traffic generation hosts called "standardHost" and "standardHost1". The traffic generated by these two hosts will be at different rates and to different destinations. Also, there are two servers to receive the sent traffic called"server1" and "server2". Hosts and servers will be connected by ethernet to routers. In this scenario, three routers ("router", "router1" and "router2") will be simulated and connected by point-to-point protocol (PPP). The router will be the source of all traffic coming from standardHost and standardHost1. Router1 will be the transit router and will be connected to a local server to receive local traffic. Regarding OTN networks, all OTN connections will be assumed as directed between the routers' links with specific bandwidth rate and delays. The rate of the connection between will vary assuming the OTN network can establish a link with any rates based on OTN ODUk-flex.

StandardHost will generate the transit traffic, which will be around 70% of total traffic. As well, standardHost1 will generate the local traffic for router1

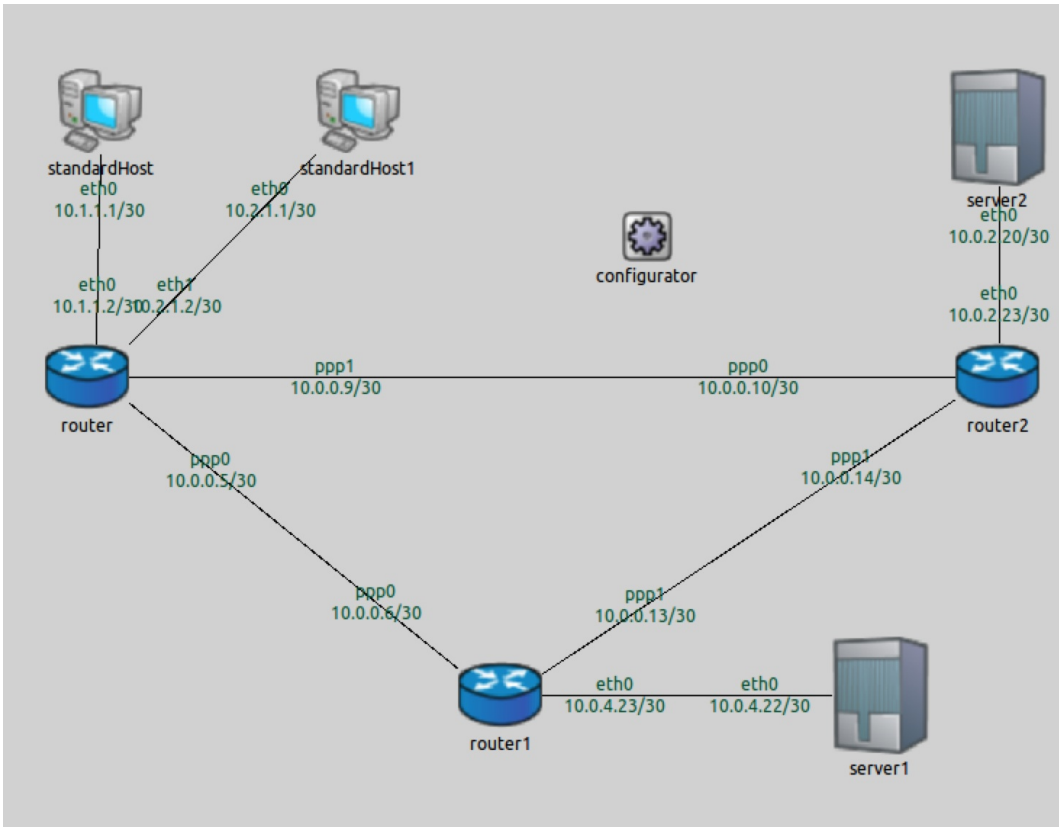


Figure 3.8: Simulation diagram (Bypassing case).

which is about 30% of total generated traffic. Different scenarios will be tested and shown in the results section.

Traffic Generation

The traffic has been generated by standardHost submodule as shown in Figure 3.8. All the traffic details are specified in the initiation file .ini, as in the following table :

	StandardHost	StandardHost1
Traffic Type	UDP	UDP
Dest. UDP port	1000	1000
Destination server	server1	server2
% of Total generated traffic	30%	70%
Packet Size	200 Bytes	200 Bytes
Intervals between packets	Random (Exponential)	Random (Exponential)

Table 3.2: Initial values of generated traffic.

All buffers are configured to handle up to 94 IP/Ethernet packets with 200 bytes per packet, which is about 170,000 bits. If any of the previous values change in the coming sections, they will be specified in that section.

3.7.3 Basic Channel Bypassing

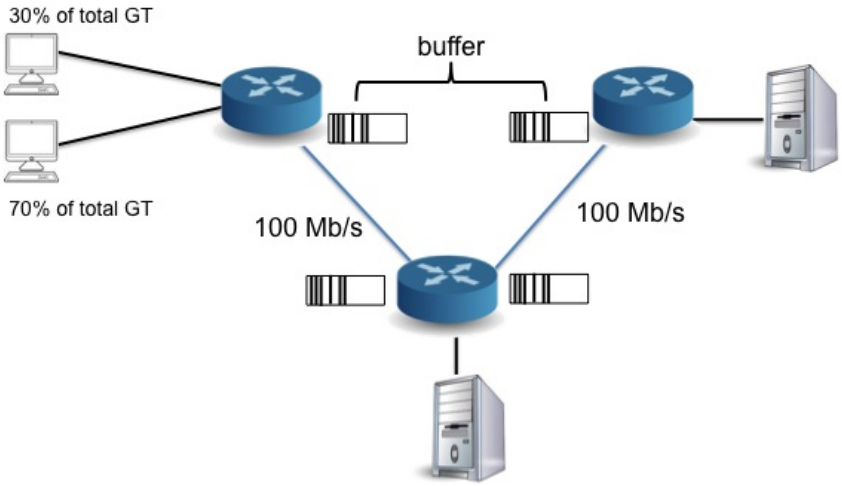


Figure 3.9: Diagram of traditional IP network (without bypassing).

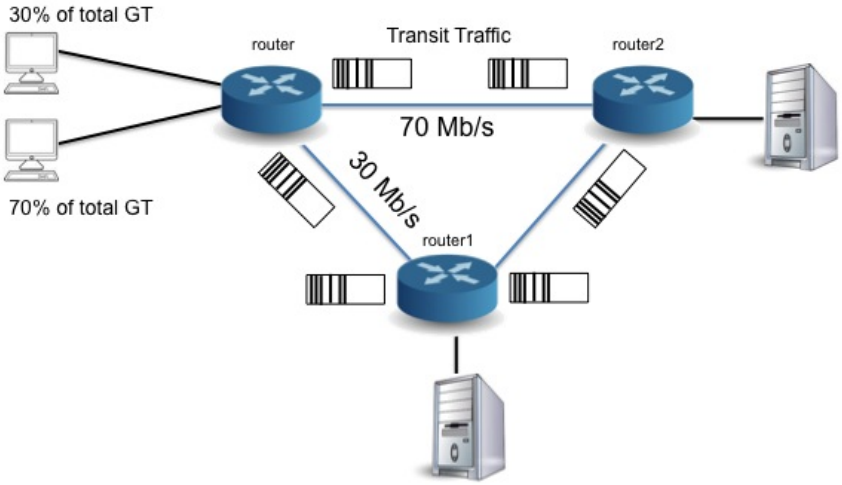


Figure 3.10: Bypassing network diagram.

Figure 3.9 shows the traditional IP network diagram. Total traffic coming from the hosts are about 100 Mbps which is the maximum rate of the WAN connection over OTN. All generated traffic will be sent to router1 including transit traffic. The transit traffic will be processed by the transit router (packet

by packet) and then sent to the next router. The local traffic will be processed and sent to the local server (about 30% of total traffic).

Figure 3.10 shows the bypassing case. In this case, there is a direct OTN link from router to router2. That link is a fixed-size link with 70 Mbps which is the amount needed to transfer the transit traffic directly to router2 by bypassing router1. In this section, statistics have been gathered to compare the

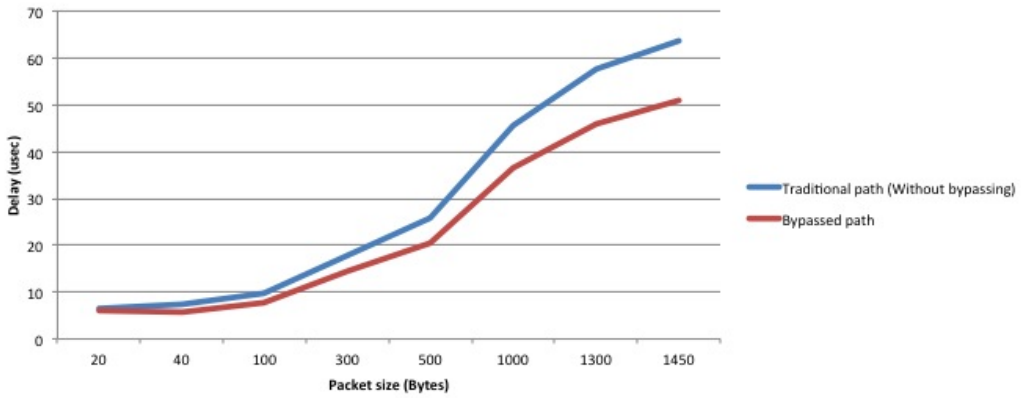


Figure 3.11: Packet size effect on delay.

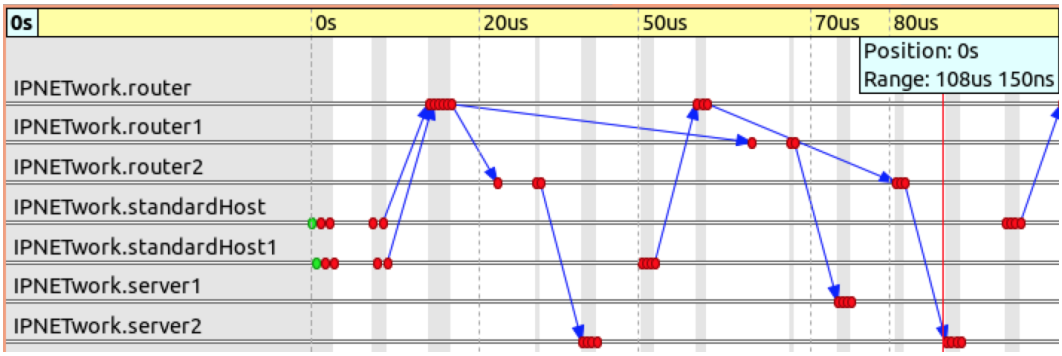


Figure 3.12: Sequential chart for bypassing path and traditional path.

two scenarios. First, we ran the two scenarios with a range of packet sizes and we noticed the end-to-end delay of the packets. In Figure 3.11, the packets are small (about 64 bytes) and there is almost no delay difference between the two paths. When the packet size gets larger (about 1450 bytes), the delay difference is about 12 μs. That delay is caused by the transmission time (processing) needed by the transit router. The main reason is that when packets are large, the transmission time will be longer. However, the delay over long distance in

core networks is in the milliseconds range, which is 100 times longer than the transmission time so it will offset any advantage of savings in transmission delays. Figure 3.12 shows packet paths going through the nodes for local traffic to router1, and transit traffic going directly to router2 by bypassing router1.

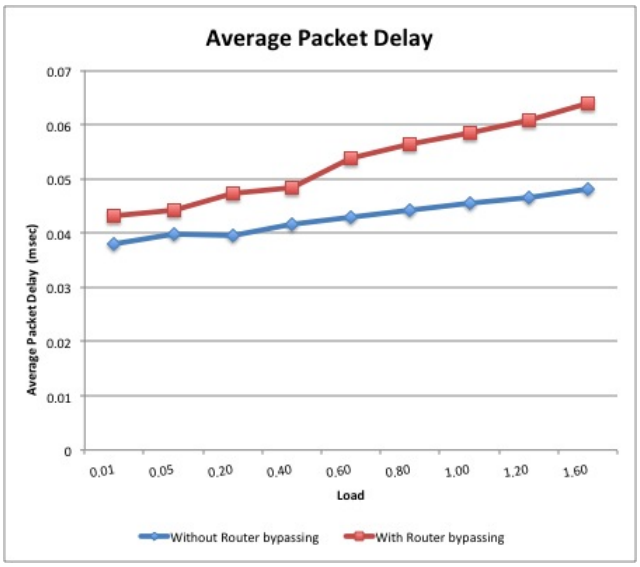


Figure 3.13: Increase of overall packet delay (in bypassing case).

As explained previously, dedicating a portion of the bandwidth to set up a direct link (circuit switched) between two nodes will lead to underutilization of the bandwidth. In our scenario, the bypassing has been set up between router and router2 to send the transit traffic (around 70%) directly to router2. Even though, at first glance, it seems that it is a faster way of transporting traffic, that is not necessarily the case. The problem is the underutilization caused by the traffic arrival probability at any router.

Traffic arrival is random with probability functions, so in our scenario at any given time probably not all 70% of the arrived packets at the router are transit traffic. In another words, at any given time, if 40% of the arrived packets at the router are local traffic for router1, then 10% of the packets need to be buffered until the link becomes available. That will increase the overall end-to-end delay of the packets as shown in Figure 3.13.

In the traditional IP network, all the traffic will be shipped to router1 regardless of its destination and router1 one will do further processing for all

traffic. Therefore, all the bandwidth will be used and no extra buffering is needed. As result, the average queuing time in the case of router bypassing

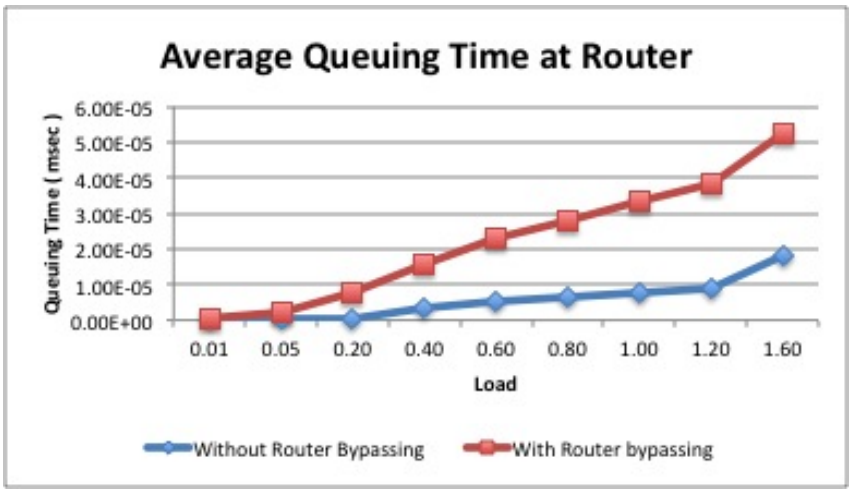


Figure 3.14: Average queuing time at router.

will be longer than the traditional way of processing all traffic by the transit routers. Figure 3.14 shows the average queuing time in our simulation for both cases at "router". In addition, the average packet loss is higher in the case of bypassing because of the increased queuing time especially when the network is experiencing high loads, a result of the increased queuing time in the routers. Every router has a finite buffer and if the traffic exceeds a certain value, it will be dropped.

In the past, these observed results have been recognized as the main downside of router bypassing. However, we assume that applying the proposed adaptive router bypassing techniques will lead to minimizing the side effects of router bypassing and maximizing the network efficiency by reducing the overall cost of OPEX (of current networks) and CAPEX (of new installed ones).

To expand that over a service provider network, Figure 3.15 shows the potential savings per number of links that are using bypassing techniques to transport traffic, assuming the transit traffic at all links is 70%. At \$0.10 per KWatts, Figure 3.16 shows how much cost could be saved in consumed power. From another perspective in IP lookup saving, which is one of the most expensive transactions inside a router for every packet, Figure 3.17 shows how many IP lookups can be saved by bypassing one router at different speeds. Since the

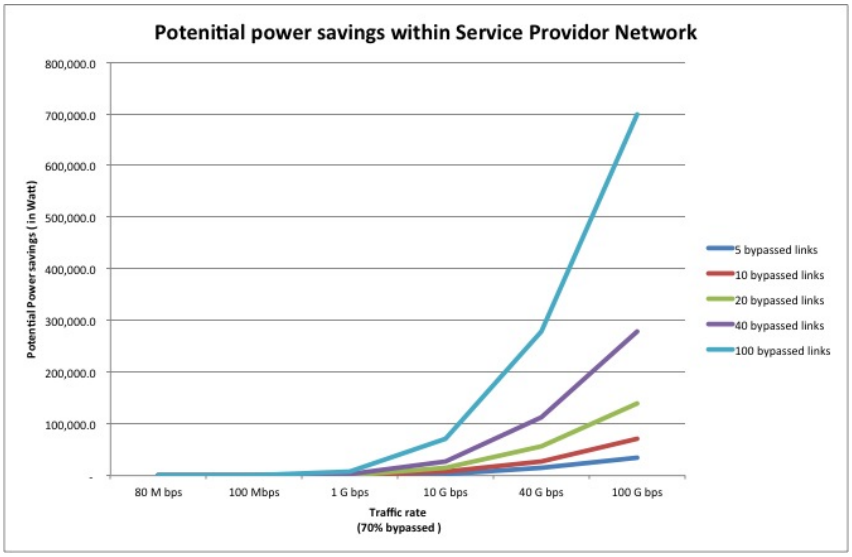


Figure 3.15: Potential power savings over a service provider network.

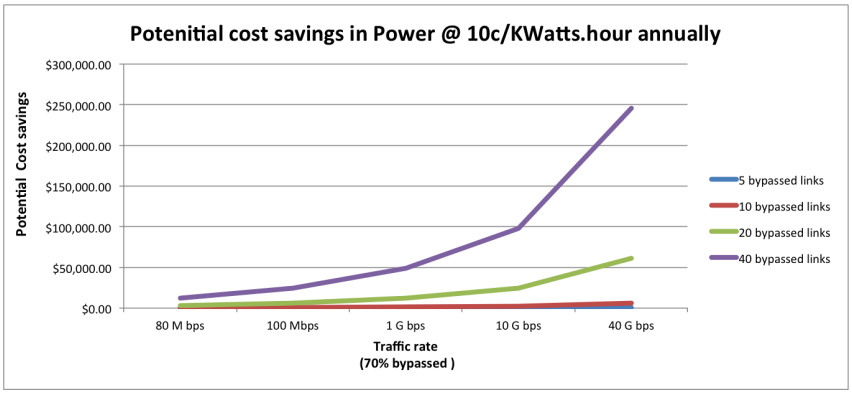


Figure 3.16: Potential cost savings in power.

packet size is related directly to the number of IP lookups for a given speed, Figure 3.17 shows the savings for two packet sizes. In terms of implementation, many details are needed to deploy this concept. Some companies have recently published [9] white papers on supporting the concept of router bypassing over OTN, but few details of the technical implementation are known since most of the developments are usually kept as internal data.

To implement the proposed ideas for an adaptive router bypassing, more complete simulation needs to be built. The goal is to build a simulation where the network is responsive to the traffic changes. When the traffic reaches a certain volume, the router bypassing link is established, and that link size will

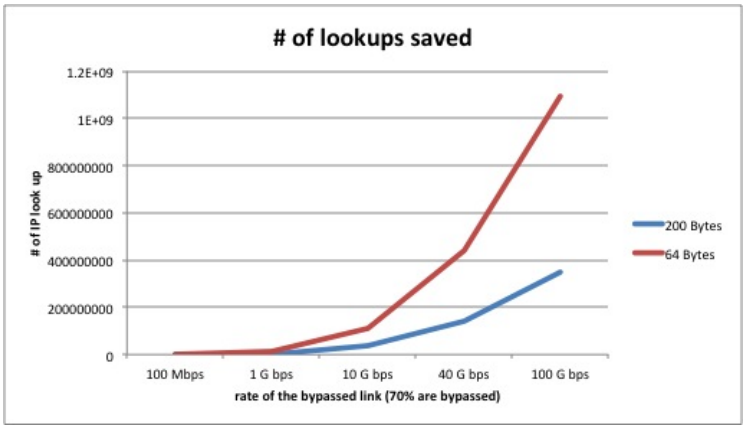


Figure 3.17: Potential number of IP lookup can be saved for different packet sizes.

be determined by traffic volume.

Statistics about traffic and traffic patterns need to be considered in this study to build the optimum router bypassing environment for the generated traffic. Traffic tends to be random with probability functions (i.e., Poisson, Pareto).

Soft bypassing needs to be implemented by creating bypassing links to minimize utilization and maximize the saving from the bypassing links. That will improve the queuing time inside the router. The degree of use of bypassing links needs to be tested and observed to get to the optimum solution.

In addition, the content-based router bypassing techniques need to be simulated. In this technique, different classes of traffic will be allocated for bypassing and the result will be observed. For instance, if smaller packet sizes use the bypassing links, then higher savings will be expected. Drawbacks also need to be observed.

3.8 Internet Traffic Models

3.8.1 Internet traffic versus business traffic

Study of the impact of time and bandwidth granularity requires a good statistical traffic model. Popular simplified models include Poisson, Pareto and

Weibull. The self-similar process is what internet traffic model follows as commonly many researchers believes [62] [63] [64]. However, business private networks and wide spread of sensors under IoT and CPS systems generates traffic from these network which does not follow necessarily the self-similar model. Gathering information constantly in a periodic fashion from sensors will lead to normal distribution or other kind of distributions such as Poisson. In this study, we assumed the network does not follow the self-similar model.

3.8.2 Aggregated Traffic

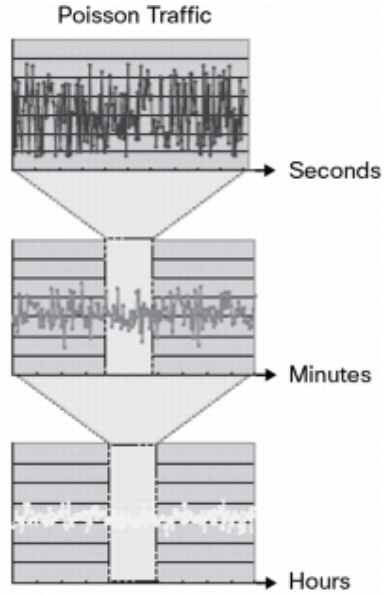


Figure 3.18: Poisson traffic behaviour over different time scales [65].

Since router bypass is implemented after traffic is aggregated on edge routers, we are interested in the traffic behaviour after aggregation. Therefore, we assume that data traffic follows a Poisson distribution. According to Cisco, the aggregated data traffic is like Figure 3.18 at different time granularities for Poisson traffic [65]. In the scale of seconds, traffic seems highly volatile and it is difficult for network equipment to adapt in response. Most network equipment requires at least a few seconds and up to a few minutes to converge [5]. However, over longer periods of time, [65] shows that traffic volume is quite close to the overall average value. Therefore, we assumed that periodic time to adjust the size of bypass channel should be large enough for the network to converge. We should emphasize that, in our proposed solution, we

assumed traffic is aggregated traffic and not from a specific user or application generated traffic.

Chapter 4

Adaptive Router Bypass Techniques

4.1 Adaptive Router Bypass using Feedback Adjusted OTN

In this proposal [66], we present a technique to enhance the performance of the router bypass by using adjustable OTN bypass channels. Bypass capacity is optimized by monitoring traffic queues and adapting the partitioning of bandwidth between bypass and non-bypass resources. The Hitless Adjust (HAO) feature in OTN is used to dynamically adjust the bandwidth allocated to bypass traffic. Since HAO enables the adjustment of the OTN channel size without service interruption, this introduces an additional degree of freedom to bypass techniques and an improvement in overall efficacy. A feedback system is used to control the adjustable bypass channel allocation. Results of a simulation (OMNET++) on a large-scale network (16 core routers, 47 access routers) show that much of the penalty incurred by partitioning the bypass capacity from the traditional packet traffic can be recovered using this adaptive bypass technique.

4.2 Background

Reduction in the achieved statistical multiplexing has been a major drawback for the bypass concept. Traditional router bypass requires subdividing

the available capacity into two fixed portions: bypass and non-bypass traffic. Traffic behaviour and volume vary over time and are not fixed. For instance, the peak 60-minute traffic volume is around 72% higher than the average daily volume [6] and the traffic growth of the peak 60 minutes is still higher than the average Internet traffic [2]. Therefore, capacity allocations should be dynamic and not fixed. Lack of visibility of the application layer (i.e., traffic volume) from a network perspective is a technological barrier for a dynamic infrastructure.

Current developments in Software Defined Networking (SDN) offer advantages over traditional networks, such as network automation, reduction in operational cost, efficient use of resources and faster deployment. SDN paves the way for an interaction between the application layer and network infrastructure. The application layer will have an abstract view of the network. That visibility will help the application layer to use the network more efficiently; therefore, network resources can adapt to the needs of the application. For instance, link provisioning can be orchestrated based on requests from an application. Application requirements such as delay, jitter, and bandwidth vary from application to application. Applications can be categorized in groups from the network perspective, based on the requested services. In a traditional network, there is no interaction between the network and application layer. This causes restriction to router bypass deployment despite its advantages. Unlike SDN, all flows are pushed into the same pipe and packet-switched by routers without any distinction between them apart from QoS. Even with traffic engineering (TE) techniques, path control for any flow is manually deployed and does not interact with traffic behaviour. In SDN, the network will transport traffic flows in an orchestrated manner.

In adaptive router bypass, we propose to offer router bypass as an SDN service for the application layer. Most studies that have explored router bypass (off-loading) have shown the advantages of this technique [8] [22] [25]. Although there are many developments in the SDN area, no other compelling study has investigated the impact of SDN on router bypass. In this paper, we explore how to enhance core networks using router bypass as an SDN service.

Traffic-based router bypass as a concept will be examined and simulated using the OMNET++ simulation tool to show the potential of SDN in router bypass. The following section will explore transport technologies that make our proposal feasible to implement.

4.2.1 Flexible Bypassing Channels Using HAO OTN

OTN has been introduced as a replacement for SONET. OTN offers more multiplexing flexibility and accepts virtually any higher-layer signals (i.e., Ethernet, IP). As one of many OTN features, Hitless Adjustment of ODUflex (GFP) (HAO) is a resizing mechanism that allows OTN to support an increase or decrease of an ODUflex (GFP) client data rate across its entire end-to-end path. HAO provides a control mechanism to increase or decrease the capacity of an ODUflex (GFP) connection without interruption to meet the bandwidth needs of the application [41].

We propose the use of the HAO feature to implement the adaptive resizing of bypass capacity. Since the OTN channel size can be adjusted without service interruption according to the HAO recommendation [41], exposing the HAO capability to bypass will allow bypass channels to be more flexible and adapt to traffic behaviour in a dynamic fashion. We assume this solution is built on top of integrated OTN/MPLS hardware such that a tight linkage exists between the control of OTN provisioning and routing. Transit traffic, for example, can be categorized based on Forwarding Equivalent Class (FEC). Each OTN switch can transport traffic to routers by default until the transit traffic load hits a certain bypassing threshold. The bypassing threshold can be set at a certain value à where the volume of transit traffic should be larger than à.

$$Vtrans > \grave{a} \tag{4.1}$$

à will depend on many factors described in the following formula :

$$\grave{a} = \beta \times C_{bp} \times \frac{T_{otn}}{T_{Router}} \times \frac{P_{otn}}{P_{Router}} \ , \tag{4.2}$$

where β: a scaling factor (dimensionless), C: total link capacity T_{otn}: time needed to reconfigure OTN switches, T_{Router}: time needed to reconfigure router (change in routing table occur) P_{otn}: power cost per bit in OTN switches and P_{Router}: power cost per bit in IP routers.

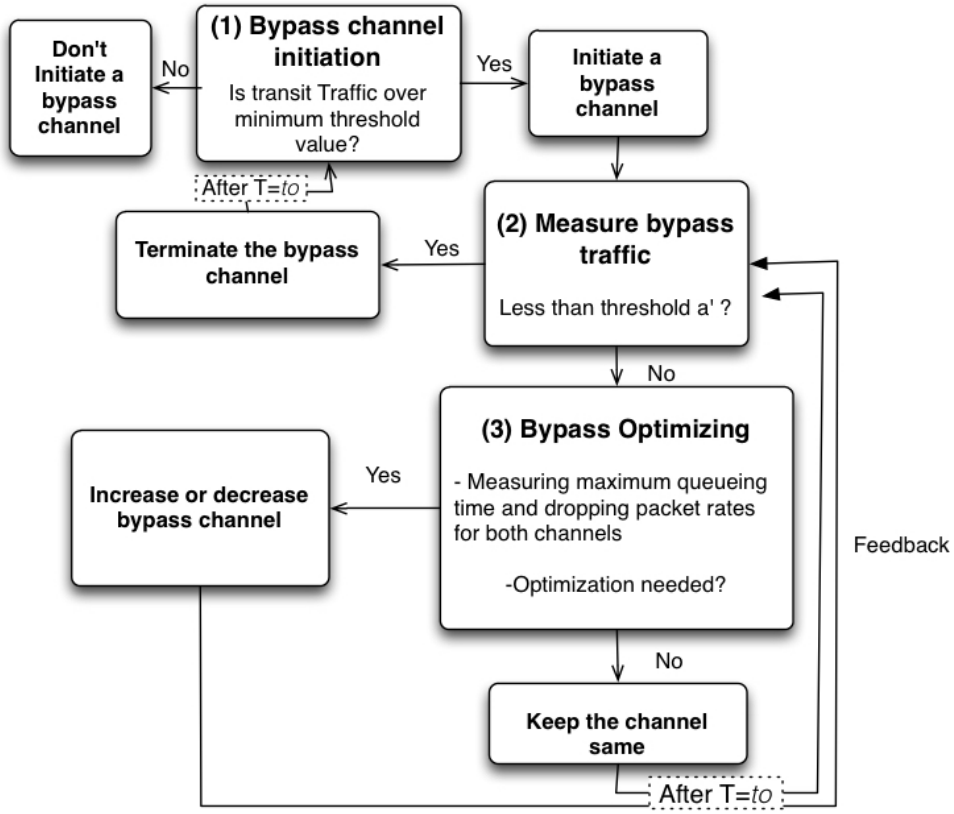


Figure 4.1: Diagram for feedback-based bypassing channel control.

Figure 4.1 is a flow chart showing bypass channel control for adaptive bypass. The bypass channel is established if transit traffic exceeds a threshold value. The amount of bandwidth allocated to bypass is then controlled by the feedback loop mediated by the Bypass Optimizer which adapts the bypass channel based on traffic behaviour. *to* is a timer to control how frequently operators review or change channel size. A smaller value for *to* means the links will be re-evaluated more frequently. That value is affected directly by the cost and complexity of the bandwidth adjustment in OTN switches. *to* has to be large enough to keep the network stable. That value will depend on technology/equipment used in the design.

In this study, optimization decisions are based on two network parameters:

(a) packet drop rate (in case of congestion), and (b) maximum queuing time for both channels. Network adaption to traffic behaviour allows focusing and using network resources where needed. For instance, during daily busy-hours, when traffic volume increases substantially, bypass be changed dynamically to handle the transit traffic more efficiently. Incremental increase (or decrease) in bypass channel size can be as fine as 1.25 Gbps (ODU0) for GFP, according to the HAO recommendation [41].

Figure 4.2 is an example where router RB is a core router and most of its processing resources are consumed by transit traffic. By using adaptive bypass, the OTNB switch will transport any transit traffic that exceeds the set threshold. All other traffic will be processed by RB. One link might have more than one transit traffic stream at the same time, as shown in Figure 4.2.

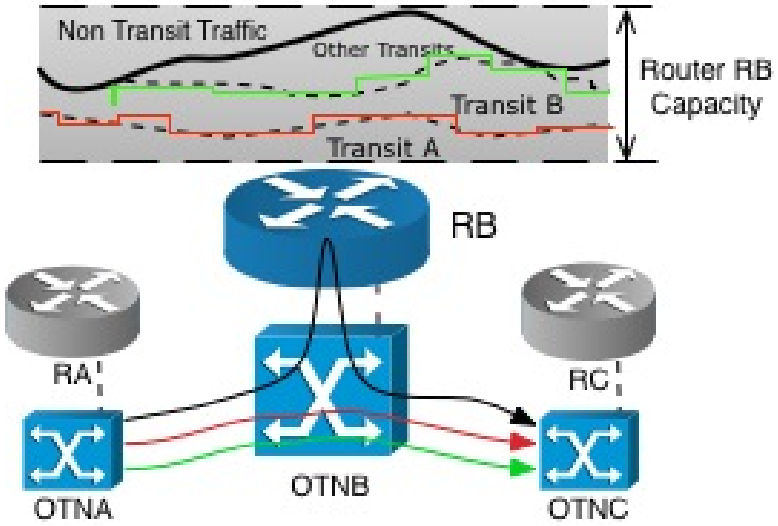


Figure 4.2: Adaptive bypassing example.

Transit channels will expand and shrink dynamically based on transit traffic volume. By applying adaptive bypass, the bypass allocation becomes aware of channel utilization and responds with fine granularity. Results shown later illustrate how this can improve efficiency.

4.2.2 Feedback-Based Utilization Optimization Technique

Traffic behaviour in the core network is volatile [42]. Therefore, allocating fixed bandwidth for transit traffic in the case of traditional bypassing is not

an efficient way to utilize the link's capacity. Moreover, adjusting bandwidth for transit traffic alone can cause a worse link throughput if other traffic is not considered. Allocated bandwidth for the bypassing channel will always be taken off the overall link capacity. Therefore, we need to consider all traffic in order to optimize the overall bypassing mechanism.

With the simplest form of bypass, one volume of traffic will be marked as *bypass traffic* and the rest as *other traffic* (non-bypass or transit traffic). In adaptive bypass, the system will monitor the behaviour of both traffic types. In this study, bypass optimization will monitor packet loss rate and maximum queuing time based on the following criteria:

Queuing Time: The optimization decision will be based on the maximum queuing time. The bandwidth for the link with maximum queuing time will be increased gradually. ϵ represents the difference in maximum queuing time.

$$\epsilon = \frac{|Q_{bypass} - Q_{otherTraffic}|}{Q_{otherTraffic} + Q_{bypass}} \ , \tag{4.3}$$

where ϵ will be minimized. The optimum case when $Q_{bypass} = Q_{otherTraffic}$ and $\epsilon = 0$.

Packet Loss: If the total traffic exceeds the link capacity, optimization will be based on packet loss rate. The bandwidth of the channel with the higher dropping rate is increased gradually until the rate of dropped packets is minimized. α represents the difference in packet loss rate.

$$\alpha = \frac{|D_{bypass} - D_{otherTraffic}|}{D_{otherTraffic} + D_{bypass}} \ , \tag{4.4}$$

where α will be minimized. If α equal zero then ϵ value will be minimized.

Loss rates and the queuing times required for optimization are obtained by monitoring queues in the appropriate hardware.

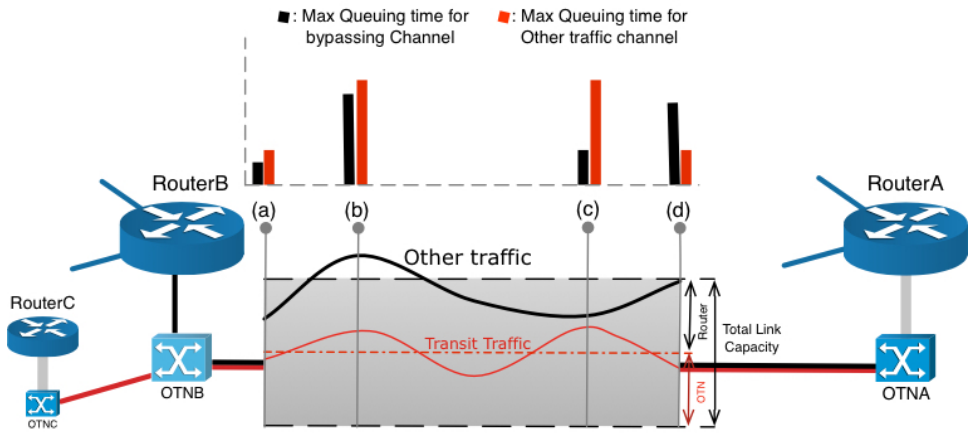


Figure 4.3: Illustration for traffic coming from RouterA going to RouterB.

Figure 4.3 is a simple a qualitative illustration for the traditional bypass scenario. RouterA is connected to RouterB through OTN. Bypass traffic will be switched by OTNB without going through RouterB. The Figure Figure 4.3 illustrates traffic from OTNA going to OTNB that comprises transit traffic and other traffic which are both changing. In this simple form of bypass, two queues will be formed: one for the bypassing traffic and one for other traffic.

Four cases have been highlighted in the example: (a) both transit and other traffic are low, (b) both transit traffic and other traffic are high, (c) transit traffic is high and other traffic is low and (d) other traffic is high and transit traffic is low. For cases (c) and (d), maximum queuing time is still high even though the total link capacity is not fully utilized. That is because in traditional bypassing, the allocated bandwidth—for the bypassing channel—is fixed, based on average volume. In our technique, by monitoring the queues in both links, the bandwidth will be adjusted to minimize the queuing time as well as dropped packets. For instance, in case (c) since transit traffic volume is high, the system will expand the transit channel because the other traffic queue is low. The frequency of bandwidth adjustments will depend on the adjustments cost in terms of time and power.

4.2.3 Simulation and Results

A simulation of a representative core network has been built using OMNET++ to test and gather statistics based on flexible OTN channels and the queue-

monitoring technique. In this case, 16 core routers (POPs) are connected to 47 access routers (ARs). The network topology is based on AT&T US IP core network [67]. Traffic is generated by traffic generators (10 clients) connected by using switches to every AR to emulate network traffic. Every source generates traffic randomly targeting all other clients with a certain rate. POPs are connected by 40 Gbps links as shown in Figure 4.4. We assume that any transit traffic switched by the OTN switch will be transported through a direct bypass link without any changes in the L3 routing table. Also, controlled traffic has been generated toward POP14 as transit traffic (going to POP2) as well as other traffic for POP6. The generated traffic is purposely controlled to produce the same highlighted points as in Figure 4.3.

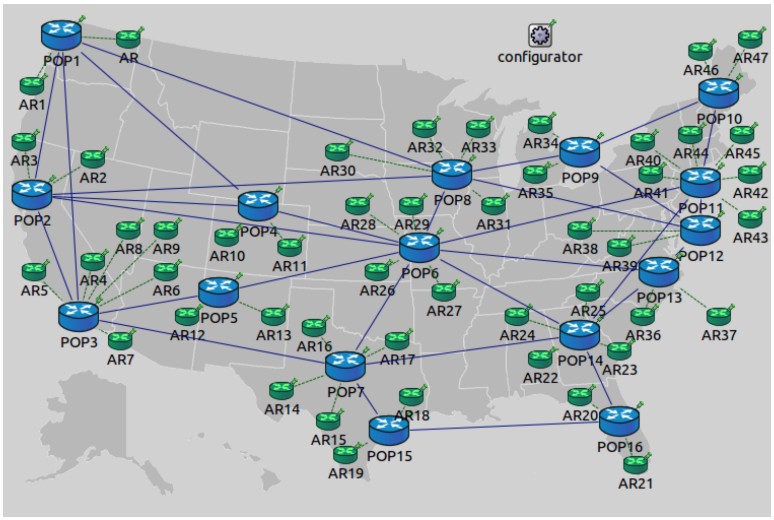
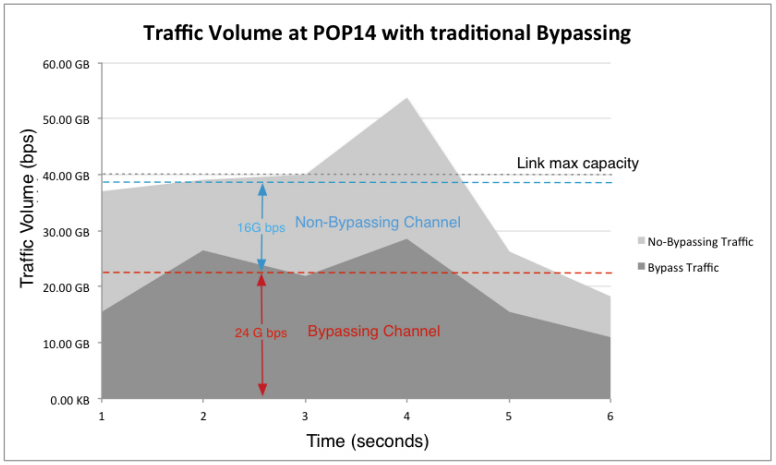


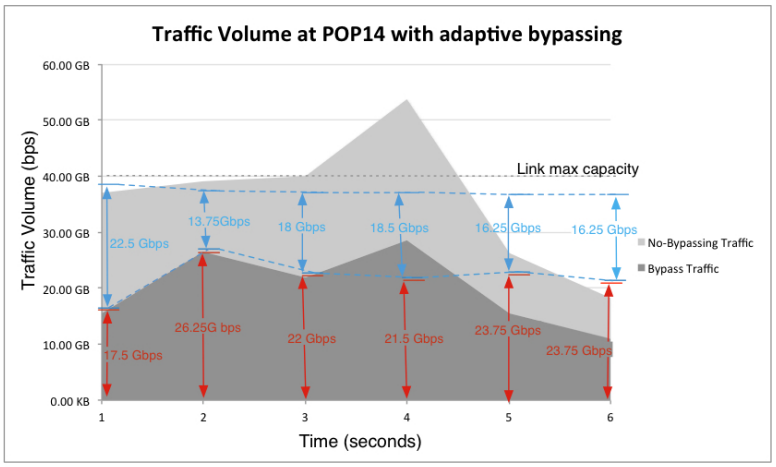
Figure 4.4: Core network simulation.

The total volume is fluctuating to capture a similar theme of Internet traffic for one day. To emulate the behaviour in peak hours, the total volume is pushed above the average by almost 40%. Figure 4.5 shows the traffic composite coming to POP14 to be sent toward POP6. Most of the traffic is transit traffic for POP6. Therefore, POP14 will initiate bypassing OTN channels to send transit traffic through OTN to POP2. Figure 4.4(a) illustrates bandwidth allocation in traditional bypassing where fixed bandwidth is allocated for bypassing based on the average value of traffic volume.Figure 4.4:(b) shows bandwidth allocation in the proposed adaptive bypassing technique. The bandwidth is adjusted over time, based on traffic changes.

In the first four seconds, the total traffic volume is almost the same as, or higher than, the link capacity. Therefore, the bandwidth is adjusted based on dropping packet rates while in the last two seconds, bandwidth allocation is optimized based on maximum queueing time. The granularity of bandwidth adjustment is ODU0 (1.25 Gbps).



(a)



(b)

Figure 4.5: Bandwidth allocation in (a) traditional bypassing (b) adaptive bypassing.

Figure 4.6 is a result for both conventional bypassing and adaptive bypassing in terms of dropping packets rate. In adaptive bypassing, the dropping rate is reduced significantly up to 66%. The reduction could be higher if the standard deviation for both traffic volumes is higher. In our simulation, the standard deviation for both types of traffic is about 7 Gbps. That is 37% and 44% of average volume for other traffic and bypassing traffic, respectively. In

the last two seconds, the total traffic is less than link capacity therefore no dropped packets occurred. The maximum queuing time is enhanced as shown

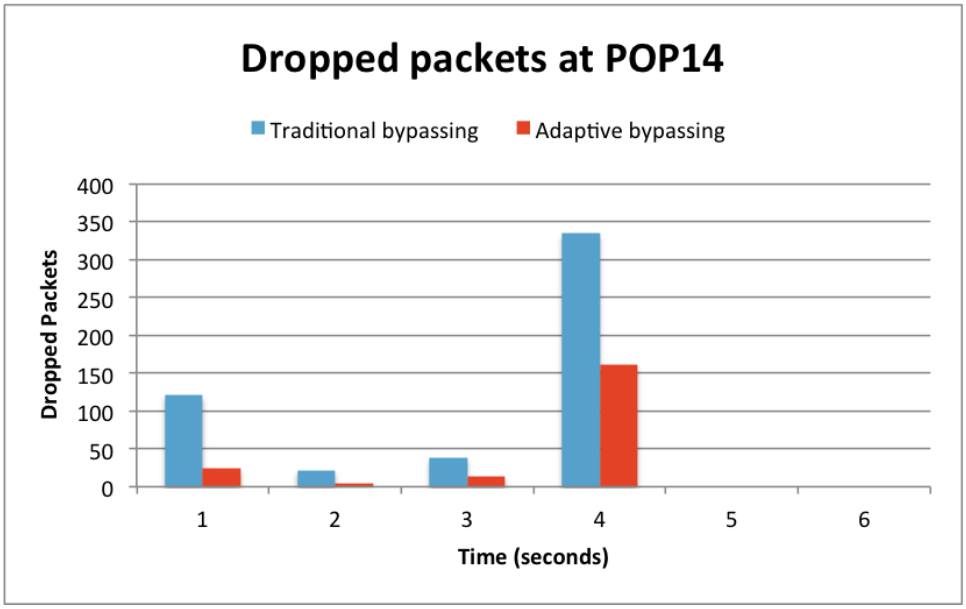


Figure 4.6: Total dropped packets in both cases.

in Figure 4.7. The maximum queuing time has been reduced over the whole simulation period.The highest reduction occurred on four(th) seconds up to 11%. As Figure 4.7 illustrates, for the first four seconds, queuing time has been reduced as well, which is predictable. However, in the last two seconds the bandwidth adjusted to purposely reduce the queuing time. Figure 4.8 shows

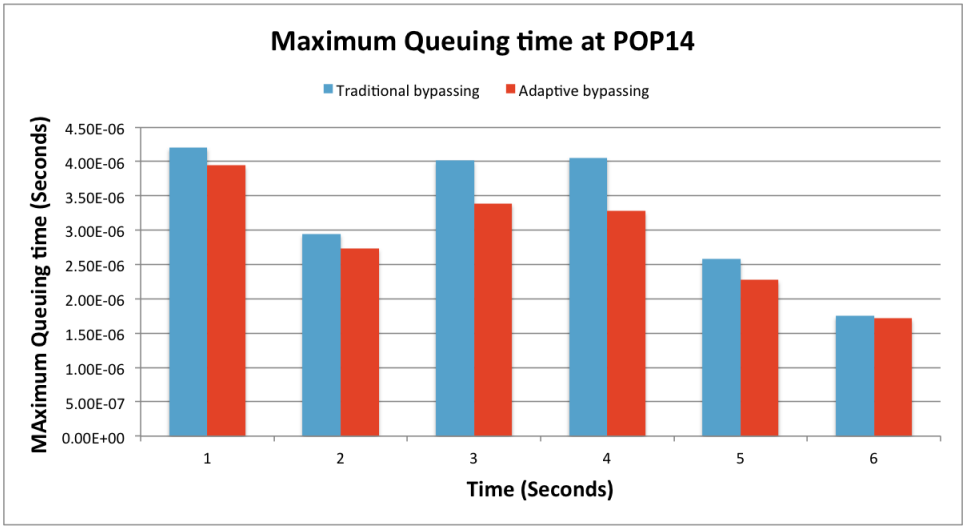


Figure 4.7: Maximum queuing time at POP14.

that, in general, throughput is increased with higher total traffic volume, as expected. There is no significant difference shown in either bypassing case except in the first second because the bypassing channel is not fully utilized where other traffic is higher than its allocated bandwidth. With bandwidth adjustment, throughput is enhanced by 6.2%.

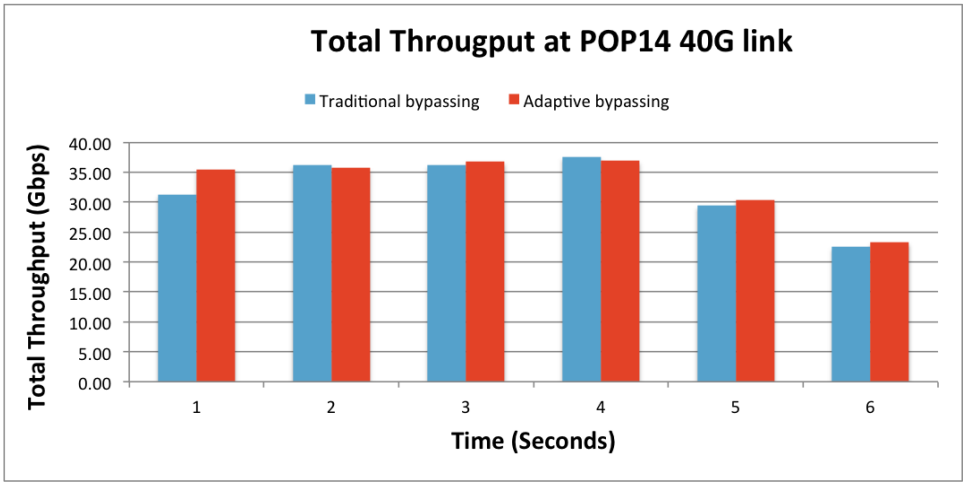


Figure 4.8: Total link throughput in both cases.

Adaptive bypassing strength comes from:

- Enhancing the statistical multiplexing issue in traditional bypassing.
- The solution is built on top of already integrated transport technology: There is no need to replace the whole hardware infrastructure.
- Maintaining tremendous savings showed in bypassing studies and deployments.

4.2.4 Conclusion

Router bypassing has shown tremendous savings in many studies of core networks. One of the main obstacles preventing the wide deployment of the router bypassing concept is a statistical multiplexing issue. In this study, we showed how new features in OTN network can optimize the bypassing implementation by introducing an adaptive bypassing technique with a feedback-based system. Adjustable bypassing channels optimized by traffic behaviour have been developed as a new adaptive bypassing technique.The concept has been

tested by a simulation of the core network built in the OMNET++ and INET framework. Adaptive bypassing has shown enhanced results over conventional bypassing. In the adaptive bypassing technique, the dropped packets rate and queuing time has been reduced by 66% and 11%, respectively.

In the adaptive bypassing technique, the reduction of statistical multiplexing occurring in traditional bypassing will be compensated for by two factors: (a) responding dynamically to traffic behaviour, and (b) adjusting the bypassing channel to have finer granularity. Future research should explore the impact of application behaviour on router bypassing to develop more intelligent and traffic-aware networks. The finer and more flexible channels offered in OTN could apply even QoS features on the transport level in a better-integrated network infrastructure.

4.3 Enhanced Router Bypass Using Fine Granularity Transport Channels

In this technique [68], we study the impact of granularity of provisioned bypass bandwidth on router bypassing. We present the use of the Hitless Adjust (HAO) feature in OTN [41] to dynamically adjust the bandwidth allocated to bypass traffic. Since HAO enables adjustment of the OTN channel size without service interruption, this introduces an additional degree of freedom to bypass and an improvement of overall efficacy. We show the potential for the improvement of router bypass efficiency by using the finer-granularity bypass scheme offered by OTN. The results of a simulation (OMNET++) setup show that much of the penalty incurred by partitioning bypass capacity from the traditional packet traffic can be recovered using finer-granularity bypassing.

4.3.1 Dynamic Bypass Using OTN

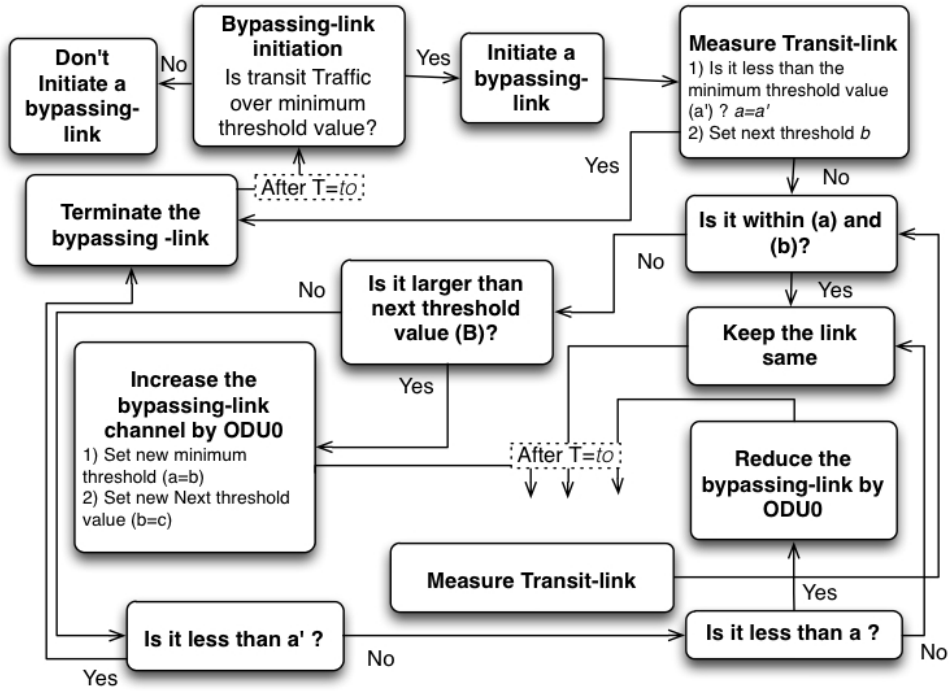


Figure 4.9: Proposal diagram for bypassing channel control.

We propose that a dynamic bypass solution be built on top of integrated OTN-MPLS hardware that offers sub-Lambda switching. Transit switching is

then implemented on a per sub-wavelength basis, not on a per-packet basis, as allowed by this integration. Transit traffic, for instance, can be categorized based on Forwarding Equivalent Class (FEC). The OTN switch will transport traffic to the router by default until the transit traffic hits a certain bypassing threshold. Instead of setting a static bandwidth for bypassing traffic (regardless of the volume), traffic volume sets the amount of bandwidth allocation needed for bypassing. This way, bandwidth allocation becomes more dynamic and responsive to the traffic demands.

The bypassing threshold can be set at initial value $\grave{a}$ where the volume of transit traffic should be larger than $\grave{a}$ in order to be switched by the OTN switch instead of the router. $\grave{a}$ will depend on many factors described in the following formula [68]:

$$V_{trans} > \grave{a} \ . \tag{4.5}$$

where $\grave{a}$ depends on many factors as described in the following representative formula :

$$\grave{a} = \beta \times C_{bp} \times \frac{T_{otn}}{T_{Router}} \times \frac{P_{otn}}{P_{Router}} \ . \tag{4.6}$$

Figure 4.9 shows the proposed scheme for channel control of dynamic router bypassing. At every state, threshold values a (lower bound) and b (upper bound) are set to increase or decrease the channel bandwidth. Channel sizes will be adjusted based on real-time data for transit traffic behaviour such as traffic volume or queuing time for both channels: bypassing and non-bypassing. For instance, if traffic volume decreases and stays below the lower bound a for a period, a new lower bound will be set and the pervious lower bound a will become the new upper bound b to adjust the channel size to a smaller capacity. In case of increases in traffic volume, the new upper bound c will be set (b=c). The advantage of using OTN is that incremental changes in channel size can be as small as 1.25 Gbps (ODU0) using GFP with HAO [44]. Time granularity will determine the value of to, a timer that determines how

frequently the channel size will be adjusted. *to* will be affected directly by the cost of adjusting the OTN channel or in case of WDM the time needed to reconfigure the WDM network.

4.3.2 Granularity Impact on Router Bypass Performance

This study shows that with an unpredictable and highly volatility traffic volume (over time scales *to* discussed above), optimizing bandwidth allocation for bypass channels over *to* combined with finer granularity bandwidth allocation can achieve better overall performance. Flexibility in OTN multiplexing allows for finer granularity bandwidth allocation that can be as small as ODU0 (1.25 Gbps). This flexibility combined with the relatively new feature that allows adjustment of the bandwidth allocation on the fly without any service interruption called the Hitless Adjustment of ODUflex (HAO) are the main motives behind this study to explore the potential of OTN in the bypass context.

In coarse bypass, increment of bandwidth change (increase or decrease) may be 10 Gbps while in fine bypass this can be as small as 1.25 Gbps, as shown in Figure Figure 4.10. Note that a WDM network takes longer to reconfigure compared to advanced OTN deployments [5]. To study the impact of bypass granularity on total bandwidth utilization, Figure 4.17 illustrates a variety of bandwidth allocation values for bypass traffic through one of the core routers. These granularities are the allowed channel sizes for OTN multiplexing [39] except the 5 and 20 Gbps granularities. Traditional bypass uses coarse bypass because of the limitations of the underlying technology (e.g., WDM, SONET). For instance, in WDM, using a single wavelength (i.e., 10 - 100 Gbps) for the bypass requires a large traffic volume to fill the capacity of an entire wavelength.

To study the impact of provisioning granularity on router bypass, bypass traffic has been generated from a source to bypass an intermediate router to reach its destination. The generated traffic followed approximately the normalized value of traffic demand reported for Google in Canada during one day [69] . Considering a 40 Gbps link as basis, the normalized data determines the

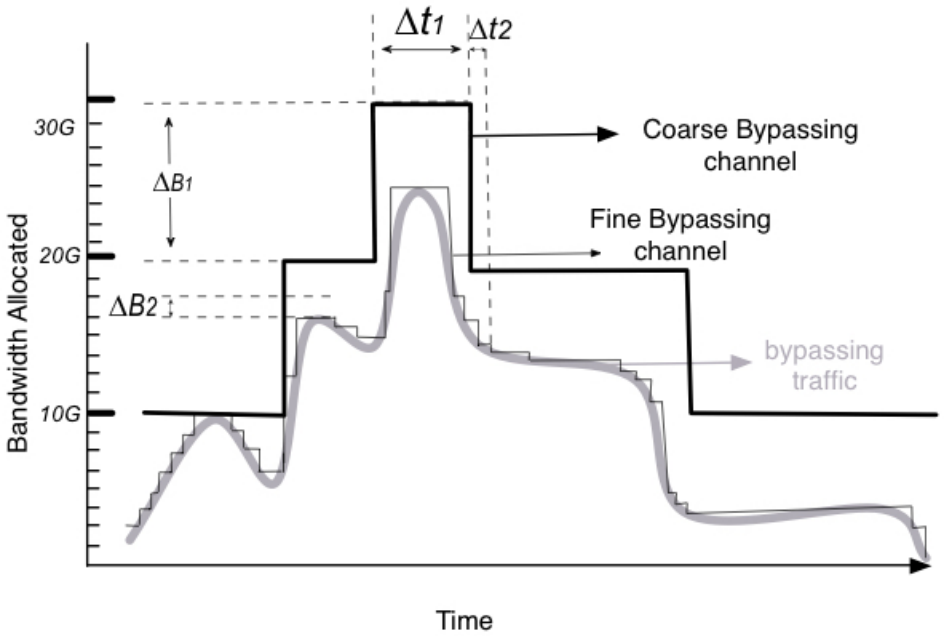


Figure 4.10: Coarse vs. finer bandwidth allocation for bypass traffic.

traffic volume with 20% standard deviation with 600 seconds accuracy. Standard deviation is introduced to emulate the unpredictability that occurs in the traffic behaviour. With fixed time granularity, the bypassing channel can be adjusted based on the traffic volume every specific period of time. In order to also observe the impact of time granularity, two sets of time granularity have been used: 30 minutes and 60 minutes. In order to determine the bypassing channel capacity, the allocated channel always attempts to accommodate all bypassing traffic. Variance granularities of channels are used to bypass traffic as shown in Figure 4.17.

To measure the excess in provisioning that occurs in all cases, we define α as the measurement ratio for over-provisioning:

$$\alpha = \frac{(C_{channel} - V_{traffic})}{V_{traffic}} \ , \tag{4.7}$$

where $C_{channel}$: allocated capacity for bypassing traffic (in bps), $V_{traffic}$: actual volume of bypassing traffic (in bps). With focus on unutilized bandwidth, the under-provision ratio is not calculated assuming that capacity will be used by other traffic.

The calculated over-provisioning ratio at different granularity is shown

in Figure 4.21. The results show that with coarser granularity, the over-provisioning is increased significantly. With 20 Gbps granularity, which is half of the link capacity, the over-provisioning can be as large as almost 50%.

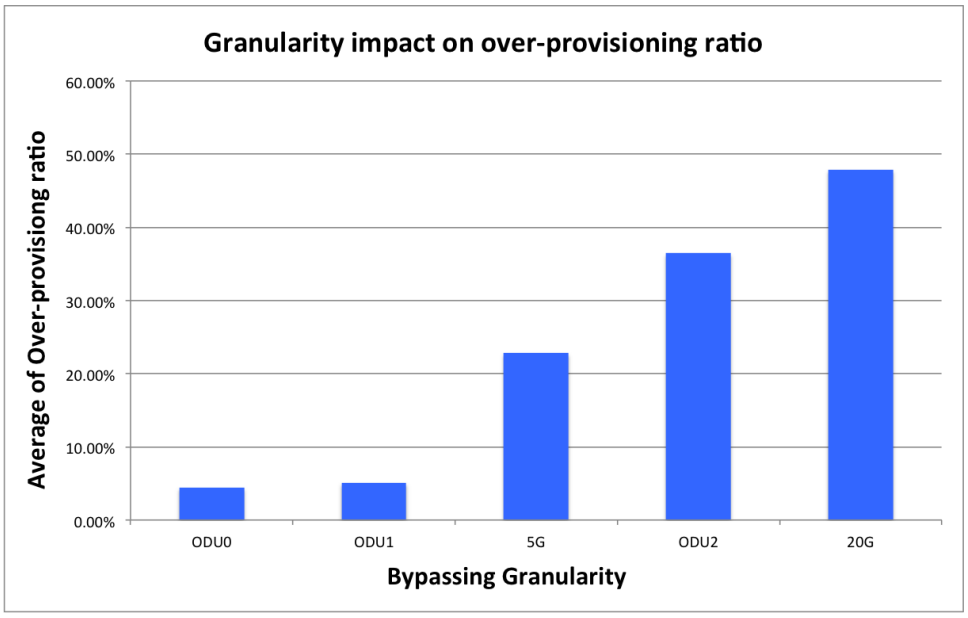


Figure 4.11: Calculated over-provisioning ratio for different bypass granularity.

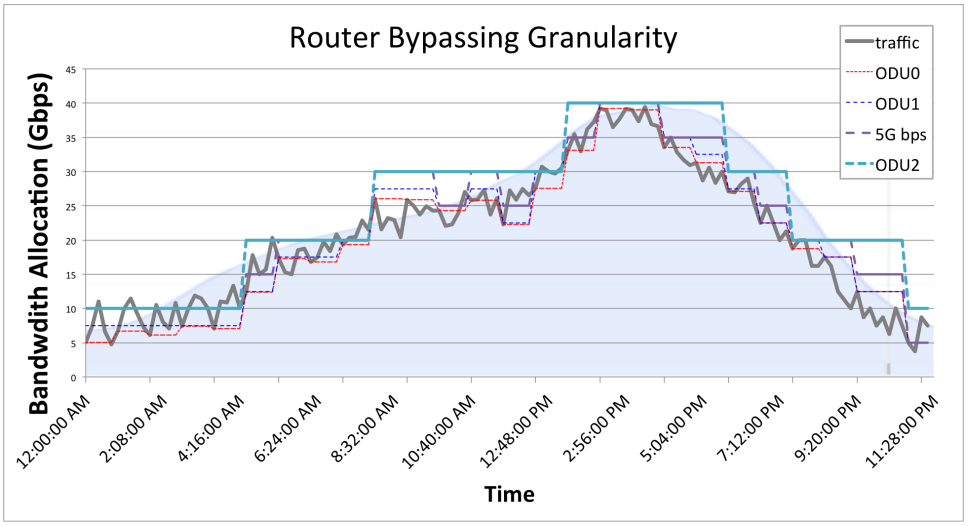


Figure 4.12: Coarse vs. finer bandwidth allocation for bypass traffic.

4.3.3 Simulation and Results

To examine the concept of bypass granularity variance, a core network simulation has been built to observe and collect statistics by using the OMNET++

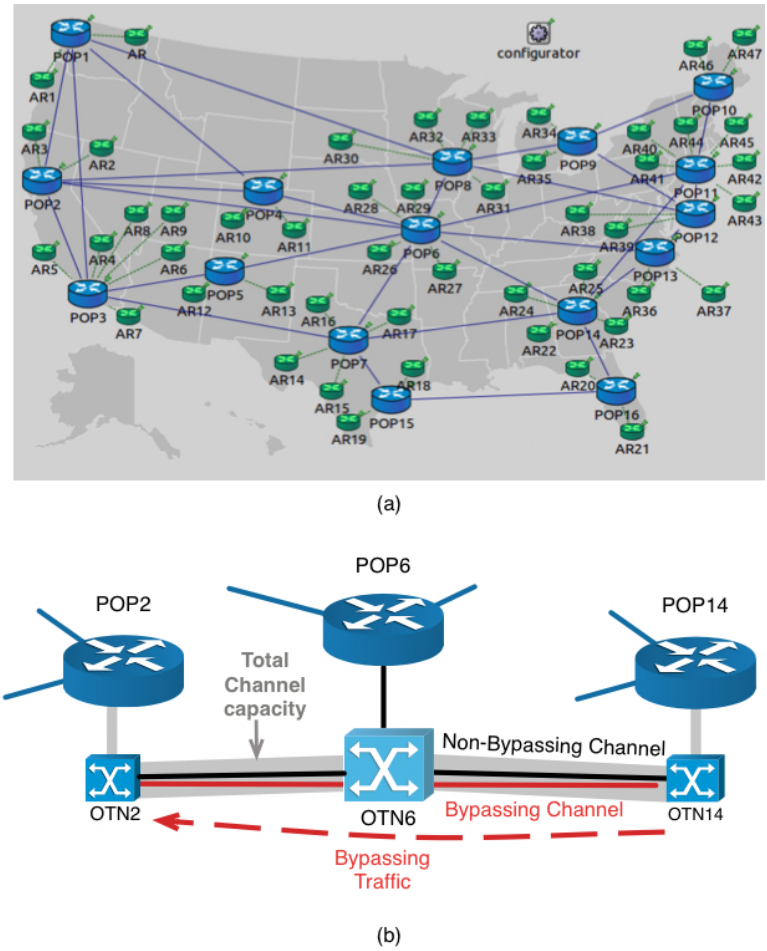


Figure 4.13: (a) Simulated core network, (b) bypass scenario.

simulator combined with the INET framework. The simulation consists of 16 core routers (POPs) that are linked together by 40 Gbps channels. Core routers are connected to 47 access routers (ARs) as shown in Figure 4.25(a). Every AR is connected to a switch serving 10 clients as traffic generators. Every client sends traffic uniformly to other clients in the network at a specific rate.

To form the bypass traffic shown in Figure 4.17, additional traffic generators are linked to POP14 to provide the bypass traffic targeting POP2. That traffic will be bypassing POP6 by partitioning the link between POP14 and POP6 as well as POP6 and POP2. The bypass link will directly connect POP14 to POP2 with an OTN connection as illustrated in Figure 4.25(b). During the simulation, the non-bypass channel is fully utilized. As a minimum, 20%

of the link is reserved for non-bypass traffic even if the bypass requires more bandwidth allocation. Both links reach congestion between 11:00 AM and 4:00 PM. The simulation has been run five times with different initial seeds; the average values are then calculated. The experiment ran on Ubuntu Linux OS inside of VirtualBox on Apple Mac Pro Platform with following specifications: 2.3 GHz Intel Core i7 and 8 GB DDR 3 memory. The experiment were stable and the standard deviation of the results is about 10%.

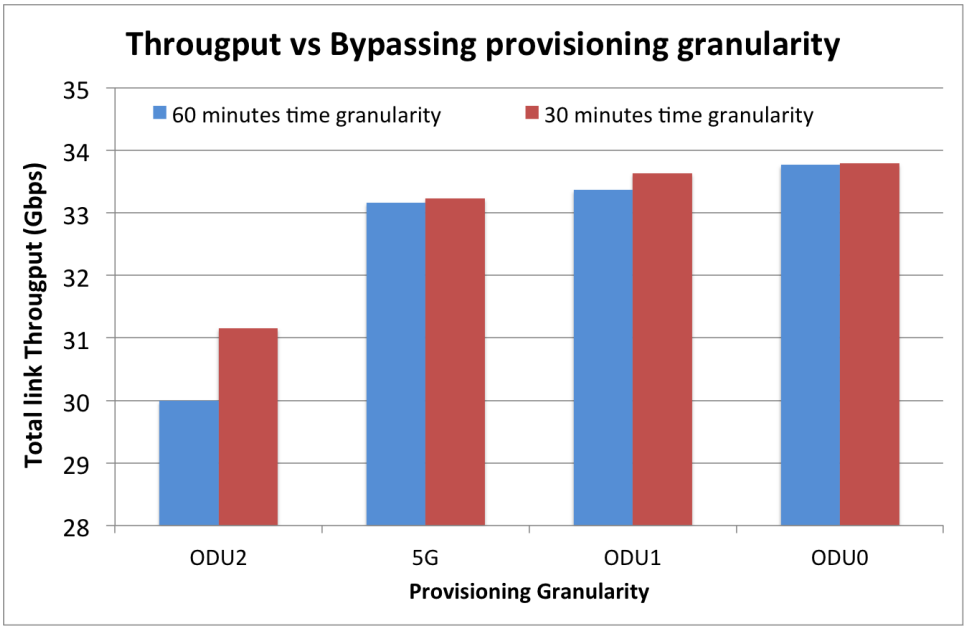


Figure 4.14: Measured throughput enhancement with finer bypass granularity.

Figure 4.14 shows that throughput is enhanced with finer bypass granularity. At ODU0 granularity (finest granularity), achieved throughput is about 8.5% higher when compared with 10 Gbps granularity (ODU2). That enhancement is because with finer granularity channels, bandwidth allocation can be very close to actual traffic volume with minimal over-provisioning. Time granularity has an impact as well; 30-minute bypass adaptation to traffic volume resulted in higher throughput when compared to 60 minutes for all bypass provisioning cases. The combined enhancement applies to both time and bandwidth granularities, and the overall throughput is enhanced by about 13%.

At peak hours, packet drop occurs because of traffic congestion. In general, with higher throughput less packet drop will occur. Figure 4.15 illustrates how finer granularity bypass reduces packet drop because of throughput en-

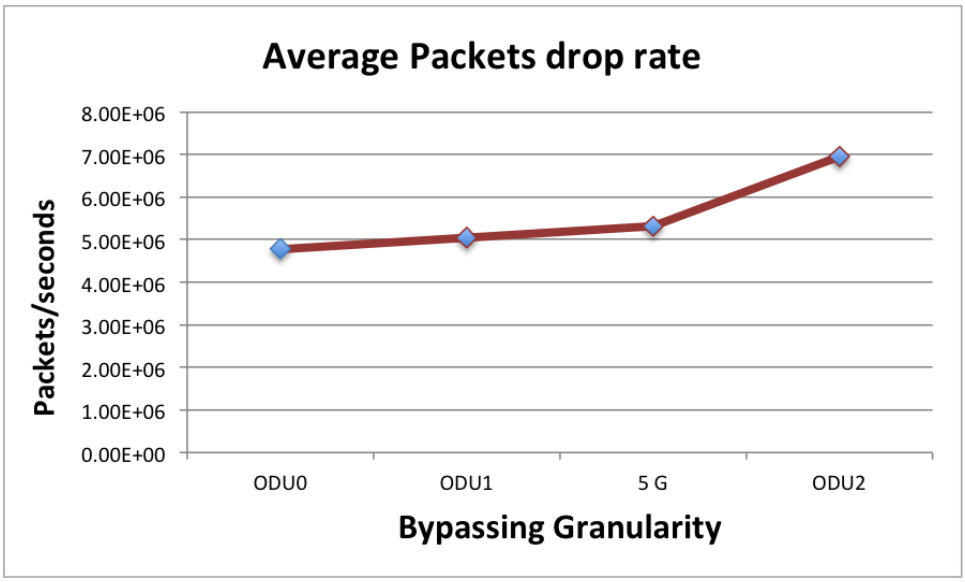


Figure 4.15: Case of congestion: packets drop rate is enhanced with finer bypass.

hancement. To simulate peak hours, bypass traffic volume from the traffic generators was increased up to the congestion state resulting in dropped packets for each granularity case. At the finest granularity ODU0, the drop rate is lower by approximately 40%. That high percent reduction occurs because we have a very congested link for which the bypass technique is used, and the enhancement caused by finer bypass will reduce the impact of congestion.

In both throughput and drop packet rate measurements, significant improvement occurred for bypass channels with bandwidth granularity less than around 5 Gbps. Even though ODU0 and ODU1 granularities have larger improvements, the correlation in incremental changes of the traffic volume—over a period *to*—to the channel granularity was the predominate factor. Hence the largest enhancement in bypassing will occur when channel granularity is close to the incremental traffic volume changes.

In summary, this study of the impact of granularity on router bypass presents the following findings:

- Coarse router bypass (e.g., > 10 Gb/s) techniques showed a higher degree of over-provisioning.

- Fine granularity router bypass (e.g., < 5 Gb/s) enhanced the overall network efficiency compared to coarse bypass.
- Fine bypass channels reduced the negative impact of bypassing on statistical multiplexing.
- Flexibility and featured multiplexing options offered in OTN paved the way for a more enhanced router bypass.
- Bypass granularity should be close to the anticipated incremental volume traffic changes over a given period.

4.3.4 Conclusion

OPEX and CAPEX savings of router bypass was shown in many studies. However, coarse bypass had a negative impact on statistical multiplexing in an unpredictable traffic environment. Advancements in OTN, such as finer granularity multiplexing, have driven this study to explore the impact of finer granularity bypassing on network performance. In this section, we measured network performance attributes, as the granularity of router bypass channels is varied based on traffic volume sampled over a certain period. The simulation experiment has been developed using the OMNET++ and the INET framework. The results showed that finer granularity bypassing enhanced the network throughput by up to 13% when compared with coarse bypassing.

With finer granularity bypass, link throughput has been enhanced by reducing capacity over-provisioning. Consequently, when the network reaches a congested state, the packet drop rate is also reduced substantially with finer bypassing. Future work will explore how to improve router bypass considering self-similar traffic and traffic shapers deployed on edge routers.

4.4 Optimizing Router Bypass Granularity Based on Traffic Behaviour

We propose that a dynamic bypass solution be built on top of an integrated OTN-MPLS hardware that offers sub-Lambda switching. Generalized Multi-

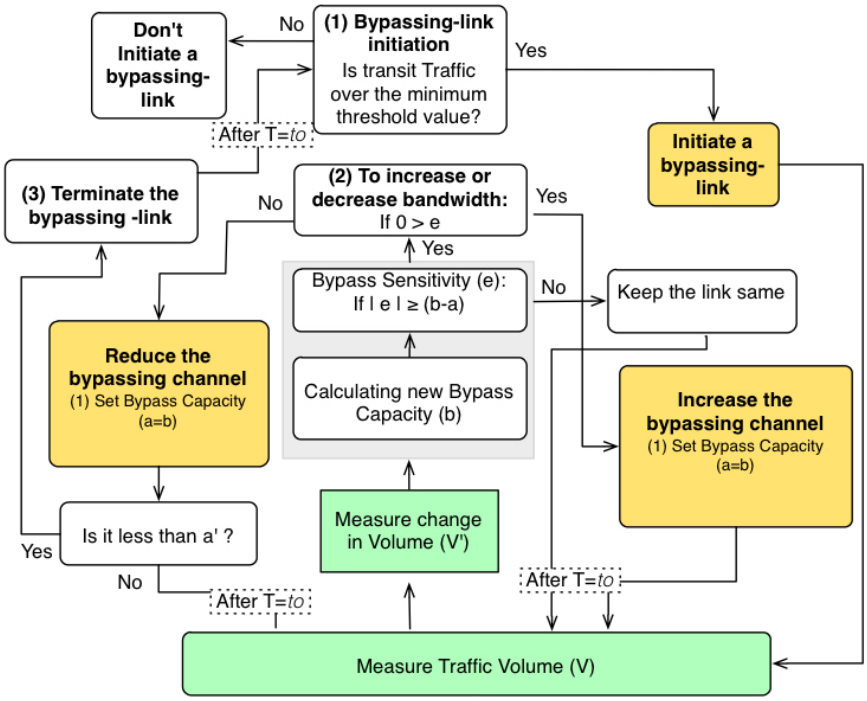


Figure 4.16: Flow chart for bypassing channel control

protocol Label Switching (GMPLS) will support MPLS-tagging on a per-channel instead of per-packet basis. Transit traffic tagging and switching are allowed through the integration of OTN and GMPLS. Transit traffic can be recognized based on Forwarding Equivalent Class (FEC). In a traditional router bypass, static bandwidth is allocated for bypassing traffic without considering changes in traffic volume.

In a bypass with dynamic granularity, our approach is to determine the granularity by considering the rate of change in traffic volume:

$$V' = \frac{\partial V}{\partial t} \approx \frac{\Delta V}{\Delta t} = \frac{V_2 - V_1}{t_2 - t_1} \ . \tag{4.8}$$

Therefore, when the V' value is high, larger granularity will be assigned for the next bypass incremental change and vice versa. As shown in Figure 4.16, the two main inputs to the system are traffic volume V and traffic volume change V' within the period of time t_o. The decision to increase or decrease the bypass channel capacity will be based on measured traffic volume V. The

size of the incremental change will be determined by V'.

After measuring both V and V', the new bandwidth capacity b will be set as follows:

$$b = c \times V \pm d, \quad \text{where} \quad d = f(V') \ , \tag{4.9}$$

$$\epsilon = b - \grave{a} \ , \tag{4.10}$$

where c is a dimensionless factor and d is the additional added or subtracted capacity. ϵ is a bypassing sensitivity factor. Bypass channel size changes when the new capacity change is larger than ϵ, otherwise the channel size will stay the same. Smaller ϵ will lead to more frequent channel adjustments.

Figure 4.16 shows the three main stages of the adaptive bypassing mechanism: initiation, adjusting and termination. Initiation and termination depend on $\grave{a}$ while adjustment depends on traffic behaviour. Bypassing capacity always follows traffic volume. The bypassing mechanism will be able to react to drastic volume changes by adding or subtracting d capacity.

The frequency of the channel adjustment will be determined by the time granularity of the system t_o. The time granularity will directly affect the cost of adjusting the OTN channel or, in case of WDM, the time needed to reconfigure the WDM network.

4.5 Aggregated Traffic Behaviour

Studying the impact of time and bandwidth granularity requires a good traffic model. Popular simplified models include Poisson, Pareto and Weibull. The Pareto distribution is applied to model self-similar packet traffic and is applicable mostly in edge networks prior to significant aggregation. Since router

bypass is implemented after traffic is aggregated on edge routers, we are interested in the traffic behaviour after traffic aggregation [65].

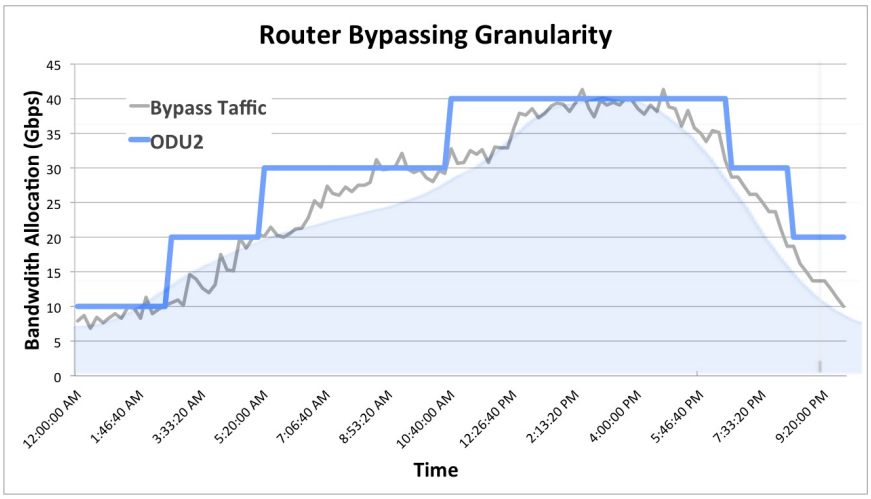


Figure 4.17: Used Internet-like traffic pattern with bypass channel with ODU2 granularity

The traffic generated in our study followed an approximated traffic pattern reported by Google in Canada during one day [69]. Considering a 40 Gbps link as the basis, the normalized data determines the traffic volume with 10% standard deviation and 600 seconds accuracy (for simplicity we denoted every 600 seconds as T). Standard deviation is introduced to emulate the unpredictability that occurs in the traffic behaviour.

Figure 4.17 illustrates the impact of bandwidth granularity on the total link utilization by a provisioned bypass channel. Bypass channel size is incremented in ODU2 (about 10 Gbps) granularity in this example and adjusted every hour.

Traditional bypass uses a coarse bandwidth bypass because of the limitations of the underlying technology (e.g., WDM, SONET). For instance in WDM, using a single wavelength (i.e., 10 - 100 Gbps) for bypass requires a large traffic volume to fill the capacity which leads to link under-utilization.

4.6 Analyzing the Impact of Bypass Granularities

With traffic volatility, the bypassing channel size can be over-provisioned or under-provisioned. In over-provisioning, the bypass channel is larger than the traffic volume and in under-provisioning the case is reversed. We need to minimize them both to maximize the efficiency of bypassing channels.

To measure the accuracy of the bypassing, off-provisioning bandwidth which includes both over-provisioning and under-provisioning is defined as:

$$\alpha = \frac{|(C_{channel} - u_t)|}{u_t} , \tag{4.11}$$

where $C_{channel}$ is the allocated capacity for bypassing traffic (in bps), and u_t is the measured volume of bypassing traffic (in bps) at point t of time. The impact of both granularity (time and bandwidth) will be explored in the next sections.

4.6.1 Time Granularity

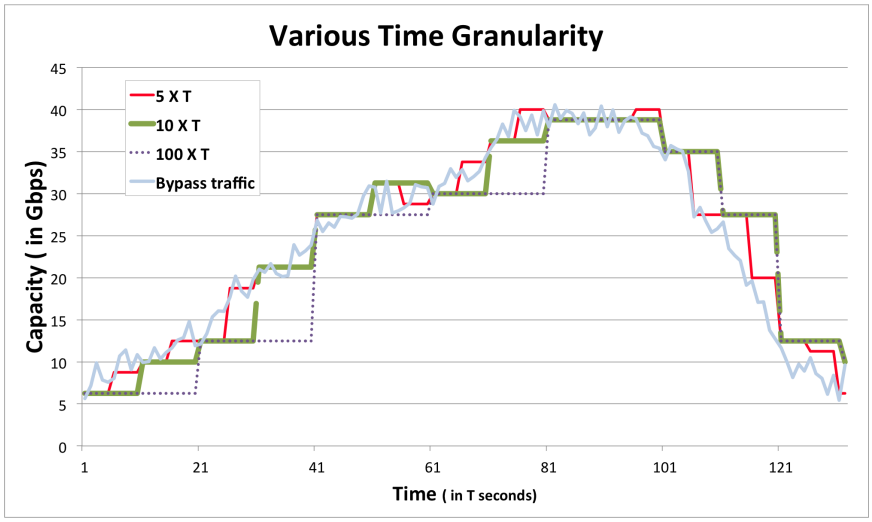


Figure 4.18: ODU0 granularity bypass channels at various time adjustments

To observe the impact of time granularity, fine granularity channels (ODU0)

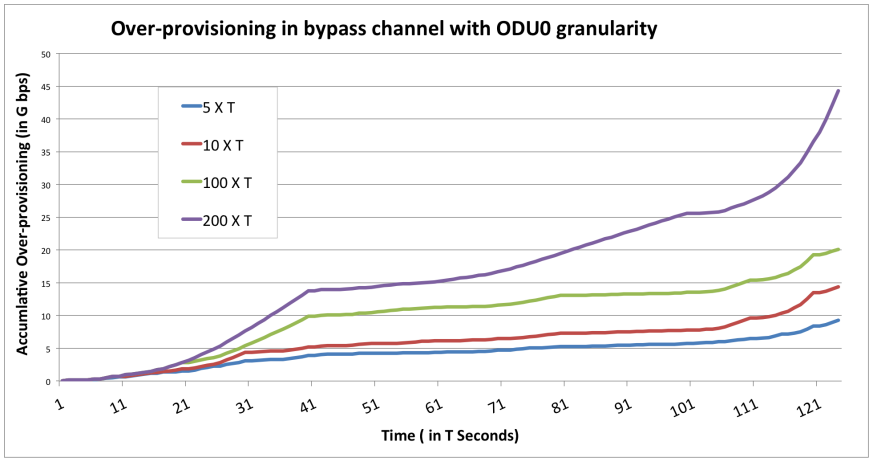


Figure 4.19: Finer time granularity provisioning

will be resized at a different time scale as shown in Figure 4.18. With traffic changing over time, channels with larger adjustment time will miss tracking traffic volumes. This will lead to higher off-provisioning as well.

Figure 4.19 shows the accumulated off-provisioning bandwidth as a result of using various time granularities, corresponding to Figure 4.18. The illustration shows that when we use finer time granularity (i.e., when we adjust the devices more often), the over-provisioning will be at a minimum. However, more frequent adjustment comes at the cost of network reconfiguration and stability. The needed device convergence time limits as to how often we can adjust the devices over a period of time. The recent advancements in OTN will achieve even faster convergence and adjusting time between the switches. In the next section, we will focus on bandwidth granularity for a given time granularity.

4.6.2 Fixed capacity granularity

With linear changes in traffic volume, fixed capacity bypass could be an optimum-option. Figure 4.20 shows an example of two fixed OTN granularities: ODU1 and ODU2. ODU1 can be assigned where the volume changes are more gradual and coarser granularities like ODU2 can be used when traffic changes are at a faster pace. With OTN, we have an option for a wide range of granularities to capture the bypass traffic. In most cases, however, Internet

traffic is not linear and tends to be very volatile during the day which is why the dynamic bypass has been introduced.

In cases where the channel's capacity can be adjusted on a real-time basis,

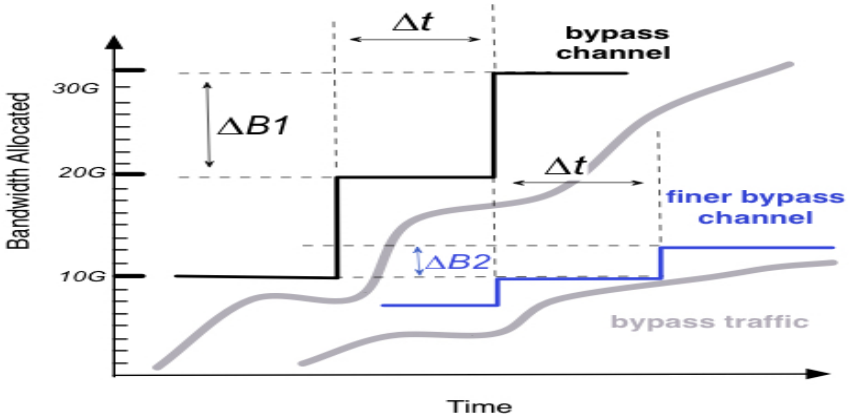


Figure 4.20: Illustration for fixed-granularity traffic bypass

(which is virtually impossible), Figure 4.21 displays that the finest granularity will continually perform better. The finest granularity will closely follow the traffic volume with minimum over-provisioning or under-provisioning. With coarse granularity, the over-provisioning can be as much as 50%.

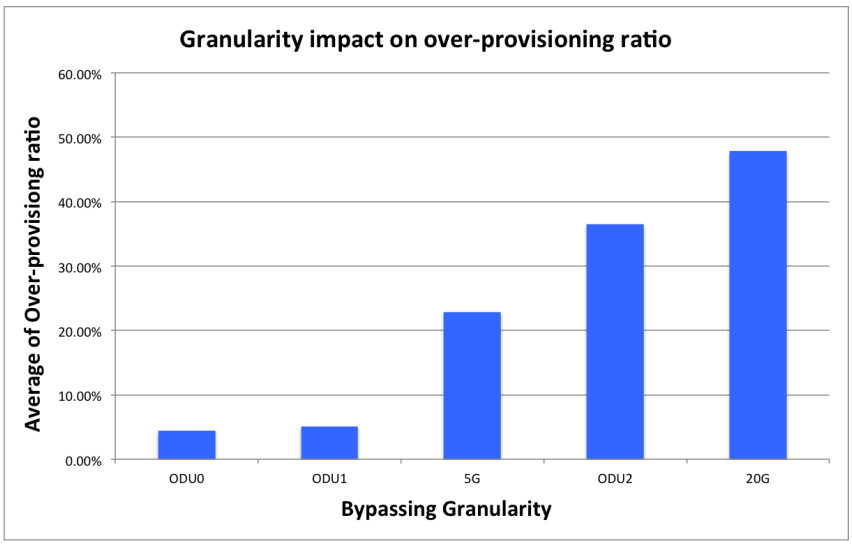


Figure 4.21: Calculated over-provisioning ratio for different bypass granularities

4.6.3 Dynamic granularity

Real-time changes in provisioning capacity is difficult or impossible to achieve; therefore, finest granularity is not always the best option if the rate of change is not linear. If traffic volume changes are linear, then fixed granularity can capture that change over time (i.e., channel size increases/decreases with traffic volume over time). However, with unpredictable traffic behaviour and with lack of linearity, any fixed-degree of granularity will not be an optimum choice.

We considered controlling the link capacity based on traffic volume changes which occurred just before the adjustment. If the calculated change in traffic within the period (differentiation) is significantly higher (or lower), then the change in capacity will be aggressive in expanding (or reducing) bypass channels. When the traffic changes in volume are small, then lower capacity will be able to capture traffic volume as illustrated in Figure 4.22.

Figure 4.22 is an illustration of both the dynamic and fixed granularity bypass. In fixed granularity, the channel granularity will accommodate the traffic volume at the adjustment time. In dynamic granularity, the traffic volume will be recorded before the adjustment time and the slope will be calculated based on equation (4.8). In Figure 4.22, we noted three points: (A) the slope of rising traffic is small and (ODU0) extra bandwidth is allocated, (B) the slope is higher and (2 X ODU1) extra bandwidth is allocated, and (C) the traffic change is more drastic and (4 X ODU1) extra bandwidth is allocated. When traffic volume decreases, capacity reduction will follow the same dynamic scheme.

When traffic surges before the rush-hour, traffic volume starts to increase drastically. With dynamic bypass, additional capacity will be added in addition to what is needed at that moment. Similarly when sudden reduction happens in traffic volume, then capacity will be reduced at a much faster rate to maximize the total link utilization. Dynamic granularity bypass, therefore, enhances bypass provisioning by observing the bandwidth change in the recent past as shown in Table 4.1. Excess capacity can be modified together with its corresponding volume change. That gives service providers another degree

Table 4.1: Example of dynamic added/subtracted capacity based on traffic changes rate

Change	Additional Capacity
20%	± (4 X ODU0)
10%-20%	± (2 X ODU0)
10%-8%	± (1 X ODU0)

of freedom to control the bypassing mechanism. Capacity changes will always be an integer number of used ODUs based on OTN specifications. Service

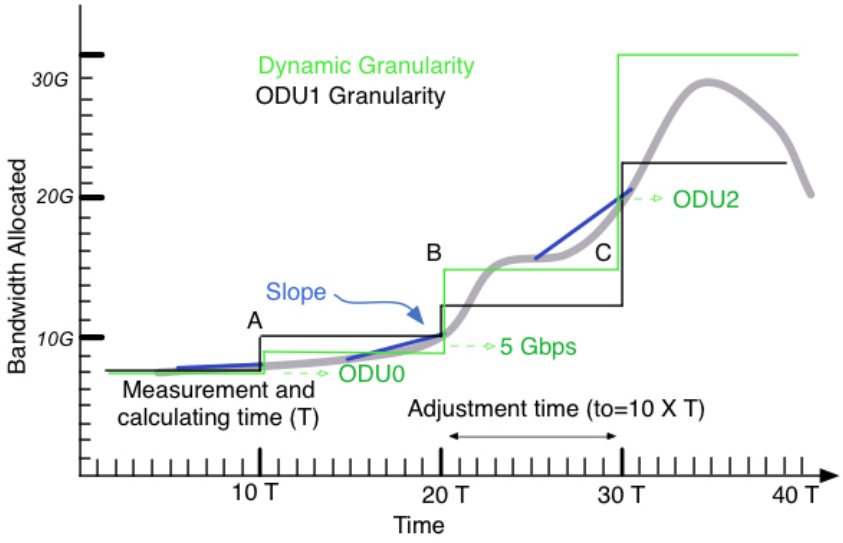


Figure 4.22: Example of Dynamic granularity bypassing: at point (A) the slope is small and 1 X ODU0 is added, at point (B) slope is larger and 4 x ODU0 or (2X ODU1) more capacity is added and at point (C) is the slope is large and 8 X ODU or (4 X ODU1) is added.

providers will be able to control a few metrics such as additional capacity, ODU size, etc. These parameters will control how aggressive dynamic bypass can be with traffic change. More aggressive bypassing leads to less link utilization which is one of the drawbacks of bypass in the first place. However, the adaptability of provisioning will improve the bypass mechanism and reduce the side effect on utilization. The goal is to maximize the efficiency of bypassing by reducing off-provisioned bandwidth as much as possible with the highest throughput. Table 4.1 shows the metrics used in our analysis.

4.6.4 Dynamic granularity and traffic volatility

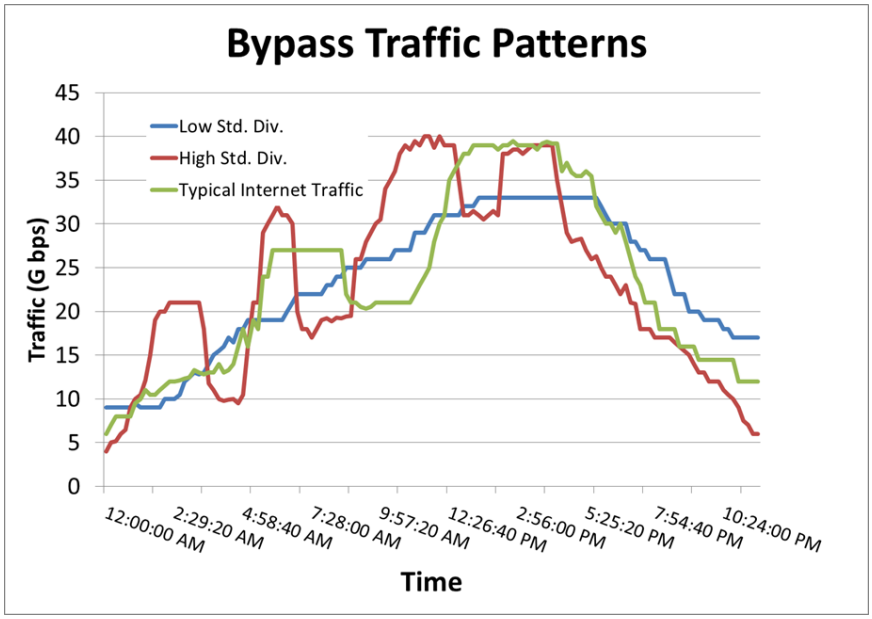


Figure 4.23: Three traffic patterns are used against dynamic bypass. All patterns have the same amount of traffic with various degrees of standard deviation: low-std, Internet and high std. The Internet-like pattern is taken from the daily Internet traffic pattern [69]

Dynamic bypass has an even greater impact when the traffic is volatile. Higher volatility requires a dynamic approach in bandwidth provisioning which results in a more efficient bypassing scheme. Figure 4.23 shows three volatile traffic patterns with the same amount of transport data.

With a more volatile pattern, traffic volume tends to surge at certain times and decay very rapidly. Typical Internet traffic surges in volume during peak-hours which is a challenge for fixed granularity bypass. In the next section, we will observe the dynamic bypass performance against these three traffic patterns. We should emphasize that there is a trade-off between these two metrics: maximizing the transit traffic capture and the link utilization. Network operators can increase the throughput by having larger bypass channels (capturing more traffic) but the off-provisioning bandwidth will be higher.

4.6.5 Analyzing the performance

To measure the performance of the dynamic bypass, the integration of the off-provisioned bandwidth α over time will be calculated as follows:

$$\sigma = \int \alpha d\alpha \cong \sum_{t=0}^{t=n} \alpha \tag{4.12}$$

The goal is to minimize the impact of the bypass on statistical multiplexing by minimizing σ. Maximizing link utilization and bypassing most of the transit traffic means σ should be minimal.

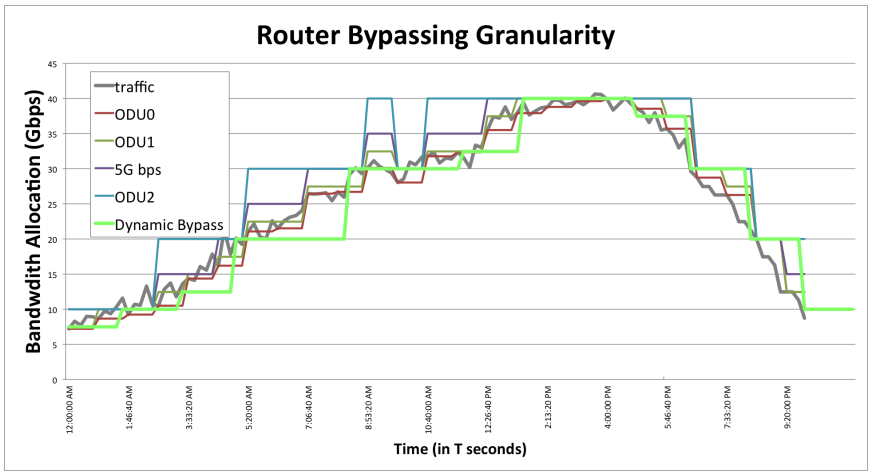
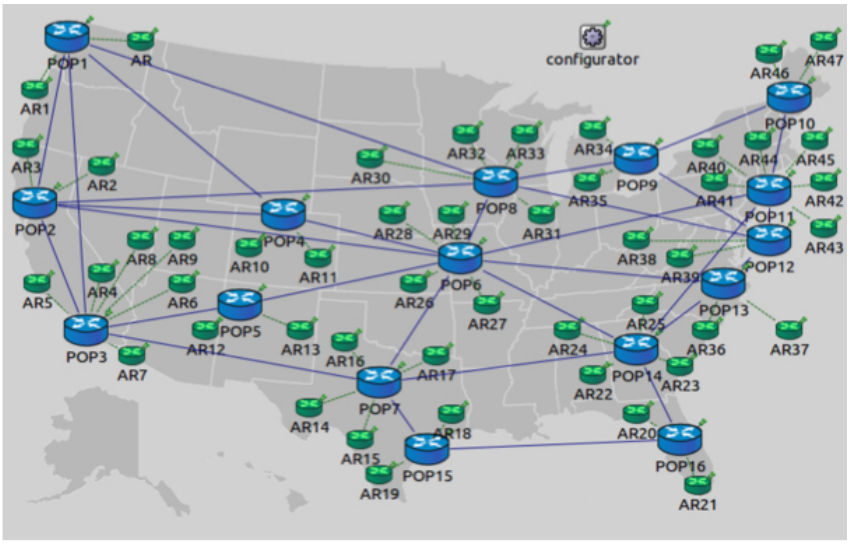


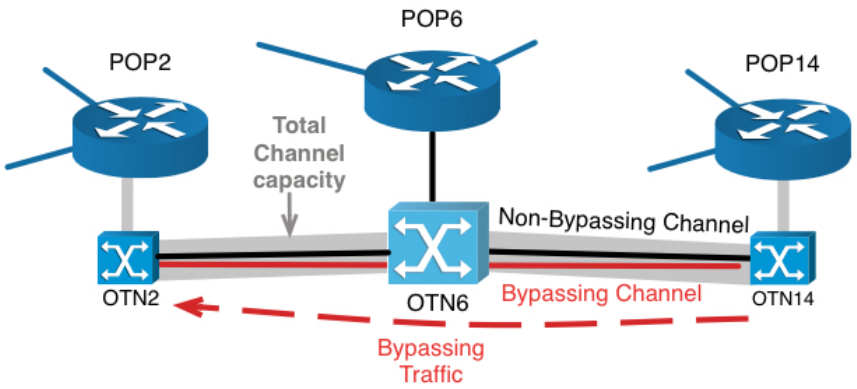
Figure 4.24: Fixed and dynamic granularity bypass channels for Internet-like traffic pattern

4.7 Simulation and Analysis

To examine the concept of bypass granularities, a core network simulation has been built to observe and collect statistics by using an OMNET++ simulator [60] combined with INET framework [61]. The simulation consists of 16 core routers (POPs) that are linked together by 40 Gbps channels. Core routers are connected to 47 access routers (ARs) as shown in Figure 4.25(a). Every AR is connected to a switch serving 10 clients as traffic generators. Every client sends traffic uniformly to other clients in the network at a specific rate.



(a)



(b)

Figure 4.25: (a) Simulated core network. (b) bypass scenario.

To simulate the bypass traffic shown in Figure 4.17, additional traffic generators are linked to POP14 to provide the bypass traffic targeting POP2. That traffic will be bypassing POP6 by partitioning the link between POP14 and POP6 as well as POP6 and POP2. The bypass link will directly connect POP14 to POP2 with an OTN connection as illustrated in Figure 4.25(b). During the simulation, the non-bypass channel is fully utilized. At a minimum, 20% of the link is reserved for non-bypass traffic even if the bypass requires more bandwidth allocation. Both links reach congestion between 11:00 AM and 4:00 PM. The simulation was run five times, enough to stabilize the results with different initial seeds and then the average values were calculated.

To observe the impact of dynamic granularity, we studied mainly two metrics: total link throughput and bandwidth off-provisioning for the three traffic patterns. The dynamic bypass approach has been examined for all three cases. Figure 4.26 shows the off-provisioned bandwidth for all three traffic patterns when different provisioning granularities are used. We can see that in all cases the dynamic granularity bypass performed better. Dynamic bypass performed significantly better with higher volatility. When traffic is unpredictable, fixed granularity provisioning will lead to more off-provisioned bandwidth and consequently less efficiency. Overall, finer granularity performed better than coarser.

All the off-provisioned bandwidth is bandwidth that could have been used

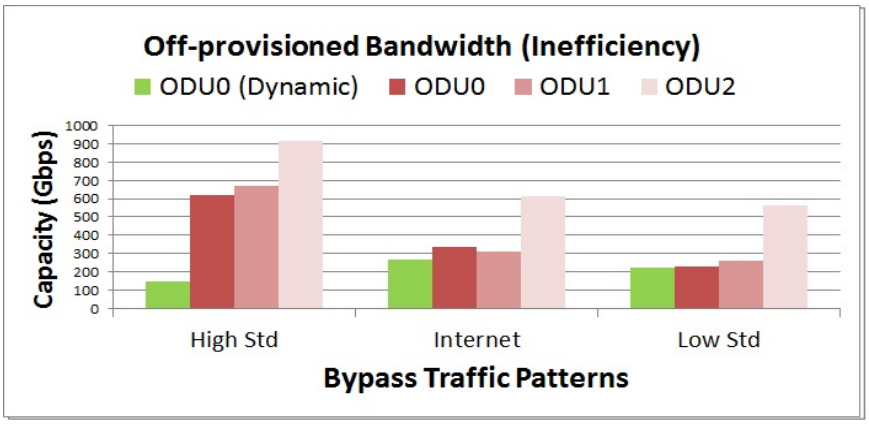


Figure 4.26: Accumulated off-provisioned capacity (inefficiency) from bypass traffic in multiple granularity. High standard deviation or volatility pattern recorded less efficiency with fixed bypass granularity. Dynamic bypass has improved the off-provisioned bandwdith results for low-std, Internet and High-std traffic patterns by 4%, 13.5% and over 300%, respectively.

to bypass transit traffic but has not because of inefficiency in bypassing which led to missing potential savings.

The amount of potential transit traffic that could have been bypassed is calculated. Figure 4.27 shows how much more traffic could have been bypassed with ODU0 dynamic bypass compared to fixed granularities bypass. We see that at fixed ODU0 bypass, the low standard deviation pattern is bypassed more efficiently than the Internet pattern. That is because ODU0 fine granularity does not allow any room for prediction to capture uncertainty or higher traffic volatility. However, with coarser fixed bypass granularities, the Internet pattern has been bypassed better. The reason is that the excess band-

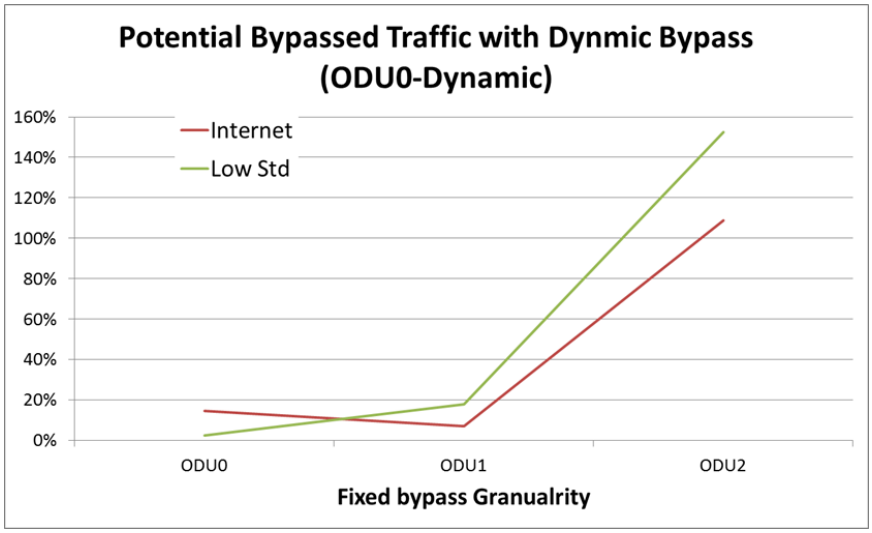


Figure 4.27: Illustration of how much more potential traffic is being bypassed with ODU0-dynamic. ODU0-dynamic bypass has more transit traffic in low std and Internet patterns by17.7% and 2.24%, respectively.

width offered by coarser granularities captures the volatility in the Internet pattern while under utilizing bandwidth. In the high standard deviation pattern, ODU0-dynamic has outperformed fixed granularities by more than 300%. These reasons show the importance of dynamic granularity bypass especially when traffic is hard to predict and is more volatile. ODU0 fixed granularity performed slightly better because the volatility is not too high.

After calculating off-provisioning bandwidth, the throughput is recorded for Internet traffic pattern to observe the performance of various granularities as shown in Figure 4.28. Similar to our previous analysis, ODU0 (finest granularity) did not perform well. Larger channel size recorded high traffic throughput simply because it captured more bypass traffic. However, dynamic granularity showed a very similar and even slightly better bypass throughout performance. That slight improvement corresponds to the improvement in off-provisioned bandwidth. Throughput can be enhanced; however, that might reduce the off-provisioned bandwidth (i.e., underutilize the total link capacity).

Network operators will be able to fine-tune how aggressively to adjust the bypass granularity with respect to traffic volume changes. A more aggressive bypass will result in more off-provisioned bandwidth but with more captured

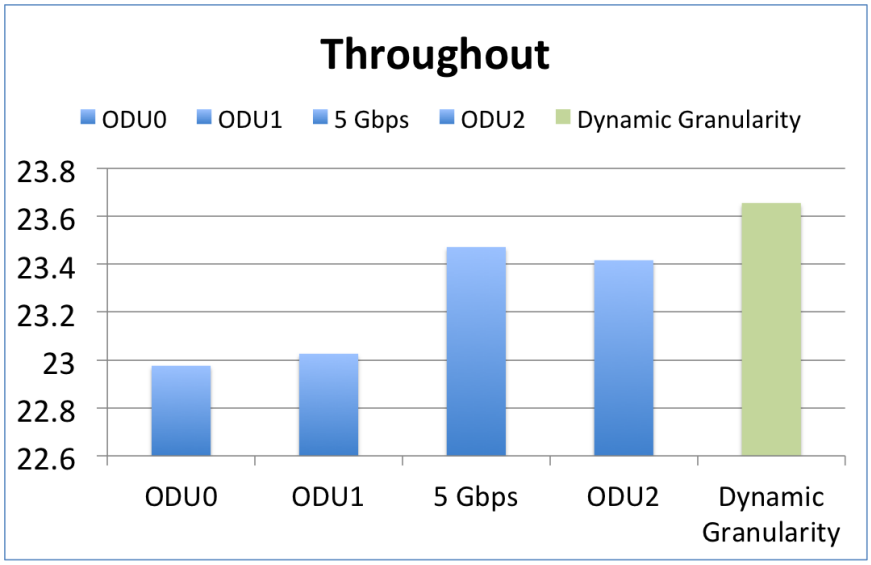


Figure 4.28: Coupled with significant provisioning improvement, throughput is enhanced with dynamic granularity channels by capturing more transit traffic

bypass bandwidth and vice versa. At peak hours, the network will be more sensitive to off-provisioned bandwidth and consequently packet drop might occur as a result of traffic congestion.

Improving the router bypass process enhances the prospect of deployment on the regional service provider's networks. Router bypass will reduce power consumption and enhance the overall switching mechanism because of the significant savings; side-effects are being minimized. Figure 4.29 shows power consumed by routers to switch traffic which may be saved through router bypass in incremental numbers of ODU1 channels based on calculations for the Cisco CRS-1 router [70]. With Internet traffic growth, these bypass savings will be tremendous in the future.

By a dynamic bypass, we allow operators to have further degree of control over the bypass process. Adding or subtracting extra bandwidth allows service providers to manage the trade-off between throughput and link underutilization in the bypass process. The goal is to minimize off-provisioned bandwidth and capture as much transit traffic as possible. In summary, this study of the impact of granularity on router bypass presents the following findings:

- Coarse router bypass (e.g. $>$ 10 Gb/s) showed a greater degree of over-

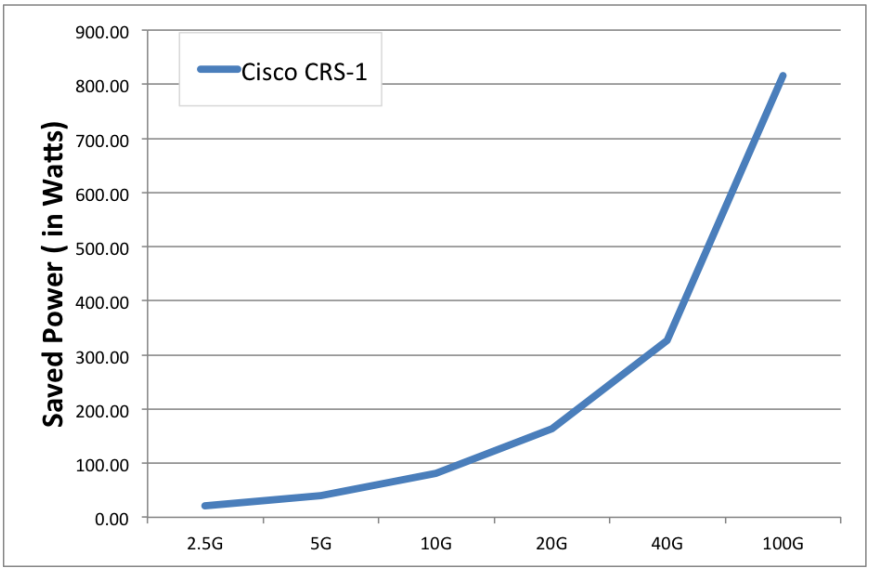


Figure 4.29: Saved power at various bypass channel sizes

provisioning.

- Dynamic bypass channels reduce the negative impact of bypassing on statistical multiplexing.
- Coupling traffic behaviour with channel size in dynamic granularity showed an enhanced overall performance.
- Dynamic bypass performs better when traffic volatility increases.
- Flexibility and featured multiplexing options offered in OTN pave the way for enhanced router bypass solutions.

4.8 Conclusion

OPEX and CAPEX savings of router bypass have been shown in many studies. Coarse bypass, however, had a negative impact on statistical multiplexing in an unpredictable traffic environment. Advances in OTN such as finer granularity multiplexing, have driven this study to explore the impact of dynamic granularity bypassing on network performance.

In this paper, we measured the network performance when the granularity of router bypass channels varies based on the traffic volume sampled over a

certain period of time. The simulation has been developed using OMNET++ and the INET framework. The results showed that when adjustment time is neglected, dynamic granularity bypassing reduced off-provisioned bandwidth by up to 300%, enhancing the network throughput slightly (by 5%) depending on traffic pattern. Considering the traffic behaviour will allow the network to react favourably. Adjusting the channels granularity as an adaptive tool is more practical than trying to predict the volume of volatile traffic. Future work will take into account channel adjustment time.

Chapter 5

Router Bypass as SDN Service

5.1 Related Work

In recent years [71], many network automation tools have been developed to facilitate network adaptability such as optimization-in-operation planning, the path computation element (PCE) and transport SDN (T-SDN). PCE can update already established connections [72] [73]. Some suggested that PCE can be used in conjunction with SDN controllers [74]. The T-SDN controller is a SDN controller to manage the transport layer using PCE according to a commercial vendor [75]. These technologies will be useful tools in the implementation of a comprehensive network automation framework. The SDN framework is widely embraced by the industry to provision efficient automated networks [76].

Rising industry adoption of SDN as an automation framework was another motivation behind this study to explore the impact of SDN on router bypass. First, we propose a mechanism of how router bypass can be offered as a network service through SDN. To observe router bypass performance, we simulate the network behaviour when the SDN mechanism is implemented. We should emphasize that this study investigates the bypassing performance with SDN traffic emulation. Therefore, the traffic-based router bypass as a concept will be generated and simulated using the OMNET++ simulation tool to show the potential of SDN in router bypass [60] [61]. The following section will explore transport technologies that make our proposal feasible to implement.

5.1.1 Optical Transport Network OTN

Dynamically adjustable channel size is needed to optimize the bypassing mechanism. Internet traffic volume changes drastically over the day (the volume at peak hours is three or four times the average traffic volume). As a result, provisioning fixed channels for changing traffic volume is not efficient. Initiating and adjusting bypassing channels based on traffic volume (demand) seems an efficient solution. OTN (G.709) has introduced a standard for adjustable channels. Hitless Adjustment of ODUflex (HAO) is a feature where OTN channels can be resized on the fly without tearing down established channels [44]. ODUflex combined with Generic Framing Procedure (GFP) channels can be resized as long as the sizes are multiples of ODU0 in the case of packet-based traffic. HAO makes OTN ready for SDN adaptability. The flexibility combined with cross-layer communication between layer 3 and the optical transport layer in the SDN framework indicate that a solid bypassing solution can be achieved.

5.1.2 Overview of Software-Defined Network SDN

The Software-Defined Network (SDN) is one of the automation solutions to handle provisioning. SDN is designed to control the network beyond the constraints of individual physical equipment (i.e., routers). Instead, controllers will be responsible for provisioning bandwidths and services in the network as shown in Figure 5.1. SDN has been defined as an umbrella of network technologies which aim to make the network flexible and agile [77]. This offers agility and flexibility and draws many network operators to embrace and even implement SDN in production networks [78]. SDN is offering a flexible bandwidth-on-demand style of provisioning which enables the network infrastructure to be dynamically customized, based on traffic demand.

The SDN framework is composed of three main layers: (a) data plane, (b) controller, and (c) application layers, as shown in Figure 5.1. The data plane contains all the physical switches and routers. These devices are governed by the controllers in the control plane through southbound APIs. In the application layer, various applications request services from the network. Communications between the application layer and controllers occur through the

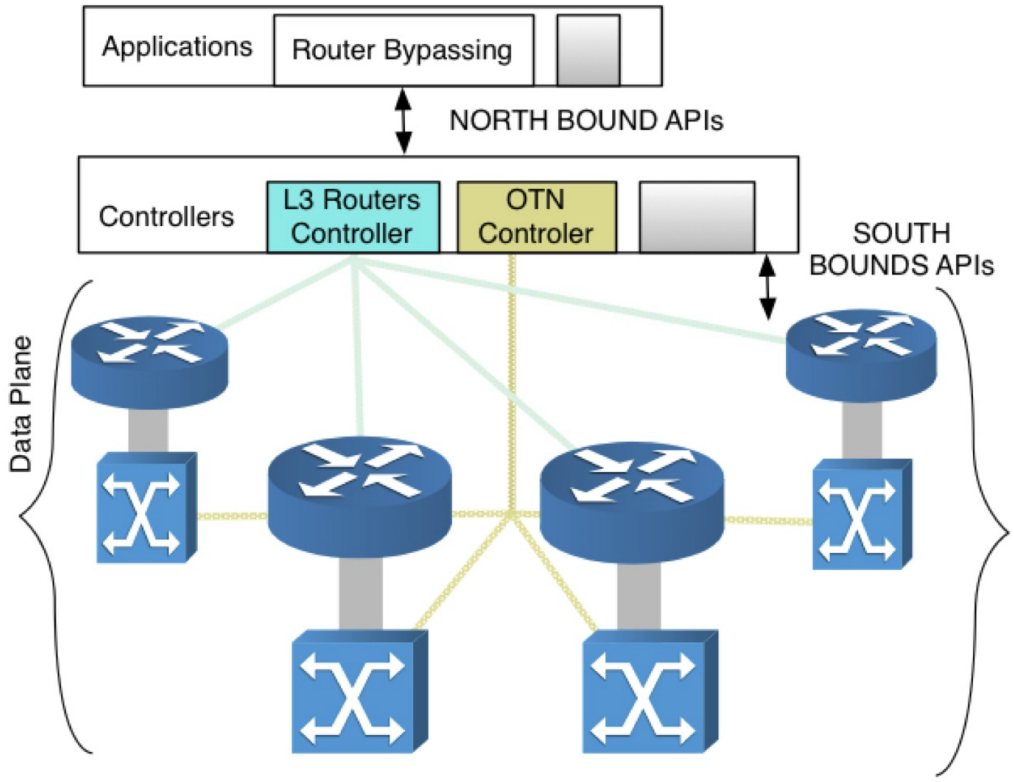


Figure 5.1: Illustration of SDN layers to control router bypass. North bound APIs allows application to request network services from controllers. South bound APIs pass reconfiguration commands to network nodes.

northbound API. Unlike the current network, applications in SDN will be able to learn more about the network through the controllers to make better decisions. For instance, application servers residing in a data centre with multiple links to the Internet will pick one link over another because the network controller informs about congestion on the other link.

5.1.3 SDN and the optical transport layer

Studies have shown that SDN-based optical networks or Software Defined Optical Networks (SDONs) can be built [79] [80] [81]. The optical controller will modify the optical equipment attributes to reconfigure switches to respond to SDN requests. For instance, building SDN over WDM using Space Division Multiplexing (SDM), where SDN an optical controller, allows control the SDM network resources by allocating certain core or spectrum to certain user or application [79]. In SDN over WDM, the SDN controller can control multiple

optical attributes in ROADMs such as a modulation scheme or the wavelength channels which will be modified based on the controller request. Other studies explored the behaviour of various optical transport technologies such as OTN and WDM with SDN. The commercial deployments of SDN over OTN is an indication of the economic incentives of SDN [82]. SDN over WDM has issues of coarse capacity assignment and the speed of the network reconfiguration [83] [84] .

Despite the economical savings in router bypass over the optical network, optical bypass over WDM faced the issue of coarse wavelength assignment [22] [85]. Without SDN even with more granular optical network technologies, the optical bypass is usually set up in static and not in dynamic fashion [5]. Our focus in this paper is the impact of SDN on router bypass as a network service regardless of the underlying optical transport technologies. However, we prefer using OTN because of its flexible granularities (i.e., ODU0, ODU1, etc.) and agility (i.e., changing link capacity on the fly using HAO without tearing down the link) which offers more flexibility to the application layer. That agility will be preferable in the SDN service-on-demand environment.

5.2 SDN for Optical Transport Network (OTN)

Advancement in OTN technology paved the way for automated and programmable optical infrastructure. Many available commercial products promise to deliver on-demand flexible bandwidth provisioning with minimum downtime [86]. As shown in Figure 5.1, the OTN controller also forms paths based on application requests. In our proposed solution, the optical controller will form physical links (i.e., connection, ODU size), which complement the logical connection of L3 controllers. Since applications initiate the service, any request should include parameters (i.e., bandwidth, time) for use by network controllers. Applications are usually aware of how much traffic will be sent to the network. Therefore, applications will request network services more economically and use network resources more efficiently. Figure 5.2 is a flowchart that shows how the application layer makes a request to the controller for a router-bypass service. In SDN, the application layer will have more informa-

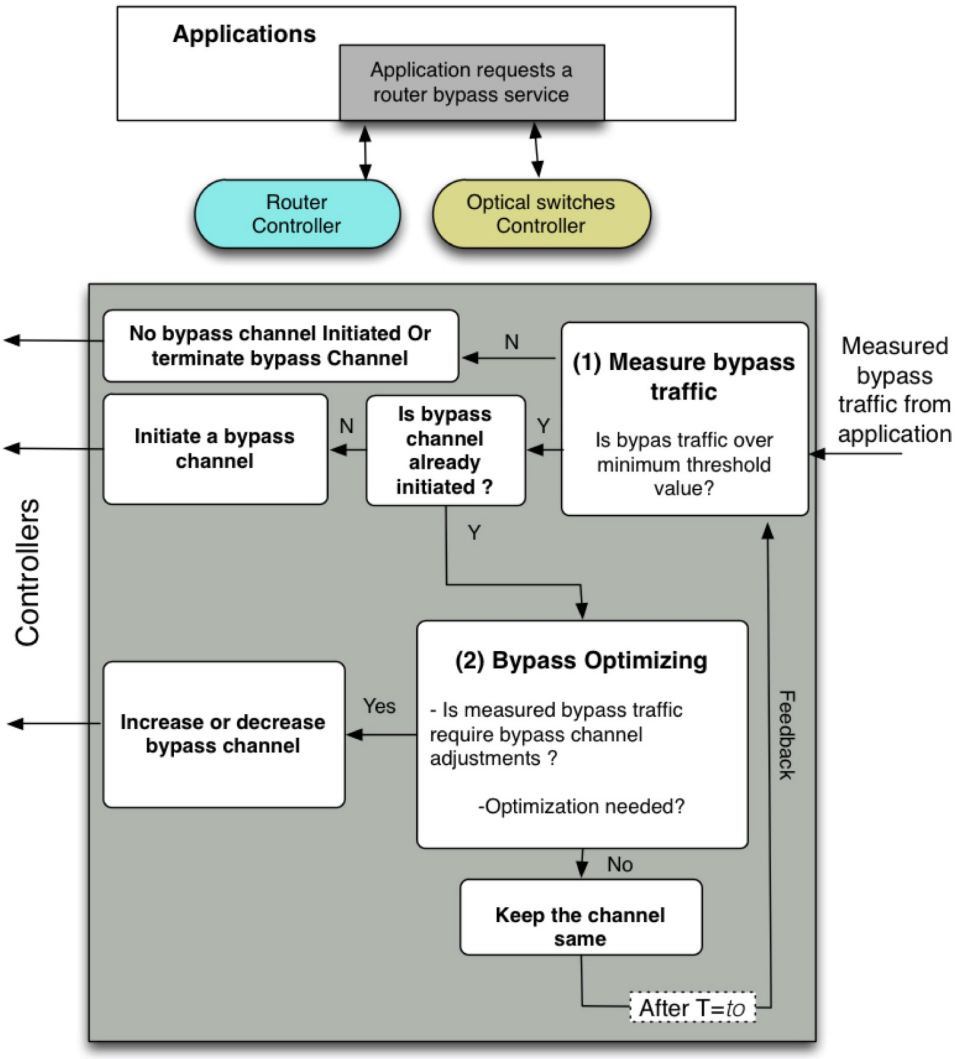


Figure 5.2: Interface between an application and controller for router bypass service.

tion and options regarding the transport network. The application layer will not only have more flexible choices, but network operators will also be able to optimize their network resources by offering more agile (flexible) options. Network circumstances might change over time and we require an adaptive interface to optimize the bypassing situation. By using traffic volume combined with a number of transit nodes, which can be obtained by the L3 controller, we can change the size of bypass channels accordingly.

5.3 Router Bypass as an SDN Service

The controller has two main roles: (a) collecting data from the networks (i.e., network topology, link utilization, congestion link), and (b) orchestrating the network based on collected data by establishing links or changing them. The central controller will have an agent in every physical device for data gathering and controlling the device by imposing any change. For instance, one approach is to use current routing protocols as an agent to gather this information (i.e., BGP-LS) [87].

Therefore, the controller will have an abstract and global view of the network in contrast to what happens in the limited network view of traditional routers. Then, controllers will provide the application layer with resource options and the application will decide how to use these resources.

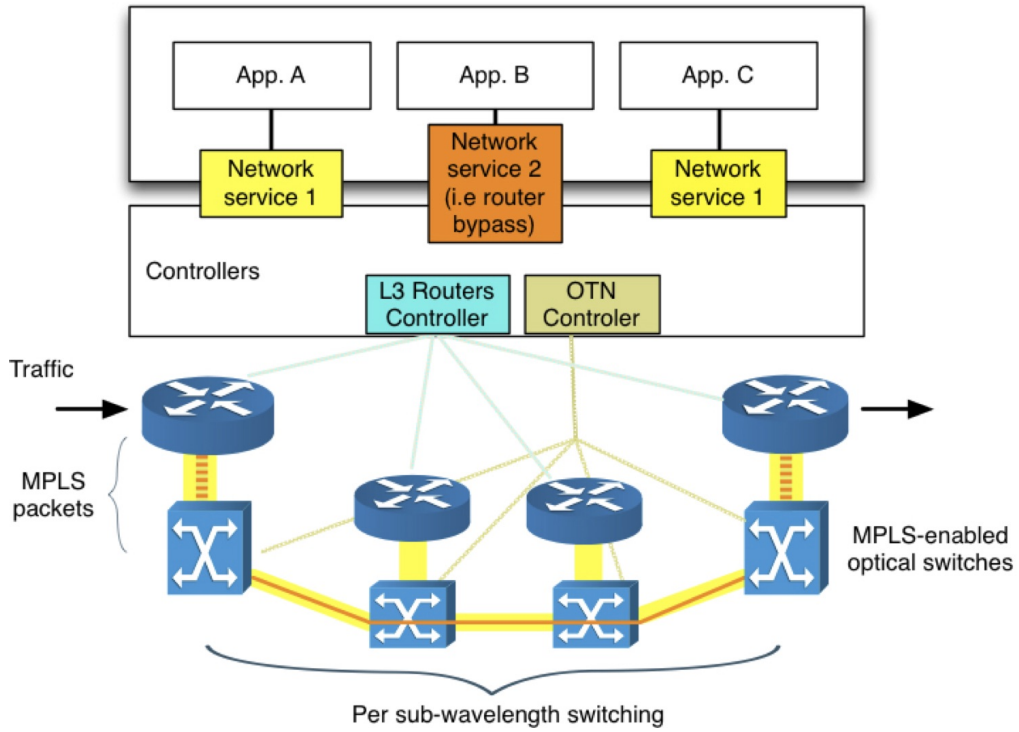


Figure 5.3: Example of router bypass service request. It shows App. B requesting bypass service from controllers. Based on that request, controllers will reconfigure the network and intiate a bypass channel.

More than one controller might co-exist in an SDN framework. Multiple controllers co-existence (cross multilayer) is described in the STRAUSS project [76]. We are suggesting two controllers in a router bypass service.

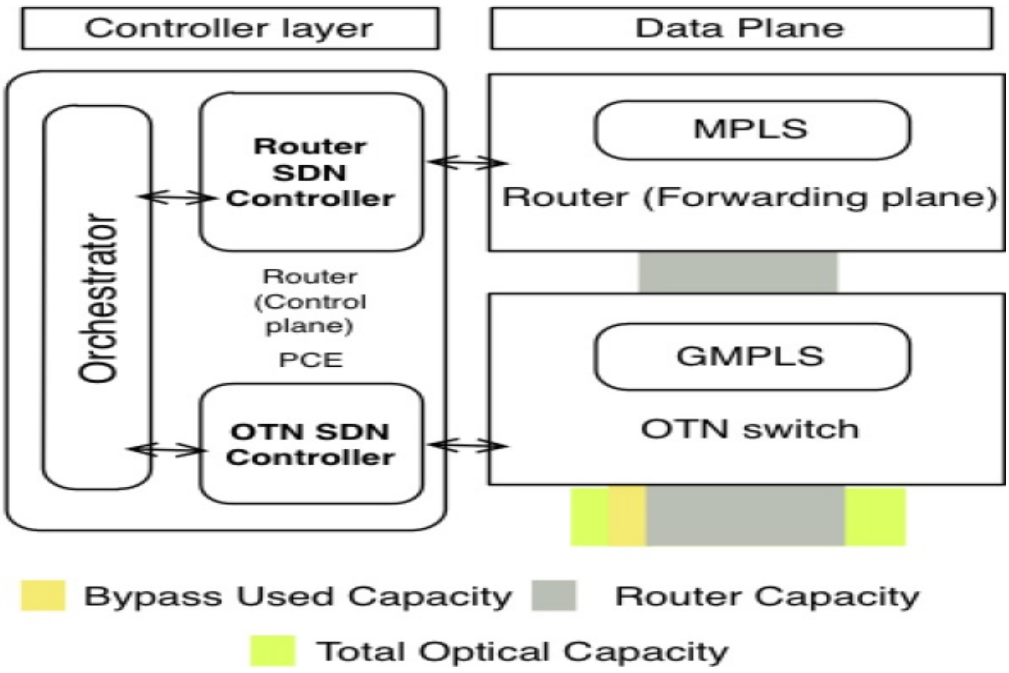


Figure 5.4: The interaction between controller layer and Data plane layer. Orchestrator will controlling how traffic will be switched: Packet switching or bypassed by OTN switch. The network intelligence exists in the controller layer including PCE to calculate the optimum bypass path and the router control plane. If the optical bandwdith capacity is larger than what the routers use, then the bypass can occur with no service impact to the local router and yet the whole bypass needs to be considered.

Figure 5.3 shows that two controllers coexist: (a) a controller for the router network, and (b) a controller for the optical network. This allows us to establish a sub-wavelength to bypass routers, which is the responsibility of the optical controller; at the same time we need the sub-wavelength to be tagged with same MPLS-tag, which is the responsibility of the router controller. The idea is to establish a path in the transport layer without changing the L3 topology (i.e., routing table). That path will be painted with MPLS colouring. The northbound request to L3 routers will trigger a change in the router's logical view of the network (i.e., change in routing table) while the request to the OTN switch will change the provisioned links or switching directions.

Every application requires a different type of network service. For instance, some applications require the least delay possible in transporting traffic, e.g., a stock-trading application. Others such as movie downloads or Torrent appli-

cations are not sensitive to delay. SDN is expected to offer services according to their requirements while using the network resources efficiently. In our proposal, we suggest that the network offers router bypass as a service (option) for applications. For example, applications that require higher bandwidth and tolerate the time needed for the network to orchestrate the service will be more suitable for a router bypass. Since the router bypass saves resources for service providers, the router-bypass service (bandwidth) can be offered at more competitive pricing than normal services. Besides the economic savings from the router bypass, traffic transported through the router-bypass service will avoid delays from the bypassed router (transmission and queuing delays). In the existing network, we believe router bypass will reduce especially the operational and Capex cost of large file transfers where millions of small packets are processed the same through the network. These can now be bypassed through the optical layer.

Figure 5.3 is an example of applications requesting a router bypass service in SDN infrastructure. Controllers offer multiple network services to the application layer. Applications request network services based on their needs. Applications A and C requested a normal service while Application B requested a router-bypass service. Next, the controller will orchestrate the router-bypass path. The optical controller will send a request for optical switches to provision the path and request MPLS-based switching on the switches. The router controller will request an MPLS tagging which will be passed to optical switches. All packets that share a common Forwarding Equivalence Class (FEC) in MPLS will be transferred through the same path. We suggest that the optical switches be Generalized Multi-Protocol Label Switching (GMPLS) compatible. GMPLS extends MPLS beyond packet switching [88]. Routers will tag the packets at the edge router and then the OTN switch will group them in a channel with an GMPLS tag. At the transit points, OTN switches will use GMPLS tagging with sub-wavelength granularity switching.

5.3.1 SDN-enabled Infrastructure Requirements

Enabling both IP and optical layers will make the network infrastructure ready to offer router bypass as service. The following requirements should be met:

- **SDN orchestrator** to supervise the optical switches and routers controllers.
- **Optical switches are GMPLS-enabled** with SDN agent to gather information and to execute the controller requests.
- **Integration in MPLS and GMPLS tagging** needed to be managed.

Exchanging information between both layers controller will facilitate that integration in the core networks. The interaction between Controller layer and Data plane layer is shown in Figure 5.4.

Dynamic bypass works better in SDN because the bypass channels will be initiated by applications, which cannot be done in a non-SDN environment. The application request for the router bypass will allow the bypassing process to be more efficient and accurate than traditional bypass. In traditional bypass even with PCE, it will be very difficult to establish an accurate initial bypass channel size for the right period. However, GMPLS and PCE will be used to switch traffic over the optical layer instead of routers. GMPLS will be used to switch traffic over ODUx (subwavelength) in OTN or wavelengths level. PCE computations will be used within the SDN controllers to compute the bypassing paths more efficiently.

A mechanism of SDN controller interaction with OTN switches has been explored [89] where application or user can be assigned to sliced bandwidth. Our focus in this paper is to observe the impact of SDN on router bypass as a concept and analyse the enhancements that can be obtained by using service-on-demand model in SDN on router bypass regardless the mechanisms used.

In this chapter, we have assumed that traffic will be processed by router (MPLS packet switching) by default. When the bypass service is requested by an application then these procedures will be followed:

- Application layer initiate request for a router bypass service with given parameters (i.e., duration of the channel, size).
- The orchestrator receives the request in the controller layer. The orchestrator should already have abstract information from controllers about

the requested path and how many routers will be bypassed.

- The orchestrator decides to initiate the bypass channel based on the collected information. GMPLS tagging will be assigned to the channel. The orchestrator will be supervising both MPLS packet tagging and GMPLS tagging.
- The application layer will receive a confirmation of the request.
- The request will be passed the OTN controller to do the required modifications.
- The controller executes the request by passing the modification to the OTN switches.
- Once the channel expires based on the application request, the dedicated bandwidth will be released and become available to the optical capacity.

The entry router will direct traffic and GMPLS tag the bypassed traffic to be switched by the optical layer. At the exit router, the traffic will be untagged and local switched to the destination.

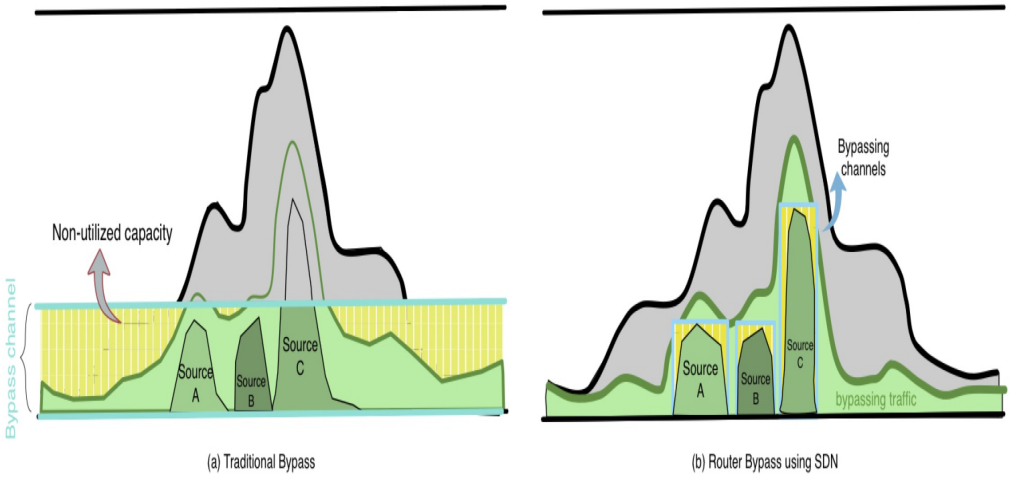


Figure 5.5: Example of SDN based router versus traditional router bypass. (a) Traditional bypass has fixed provision regardless of traffic behaviour. (b) Using SDN, bandwidth provisioning will be adaptive based on traffic behaviour.

5.4 Network Capacity and Router Bypass

5.4.1 Traditional Bypass and SDN-Based Router Bypass

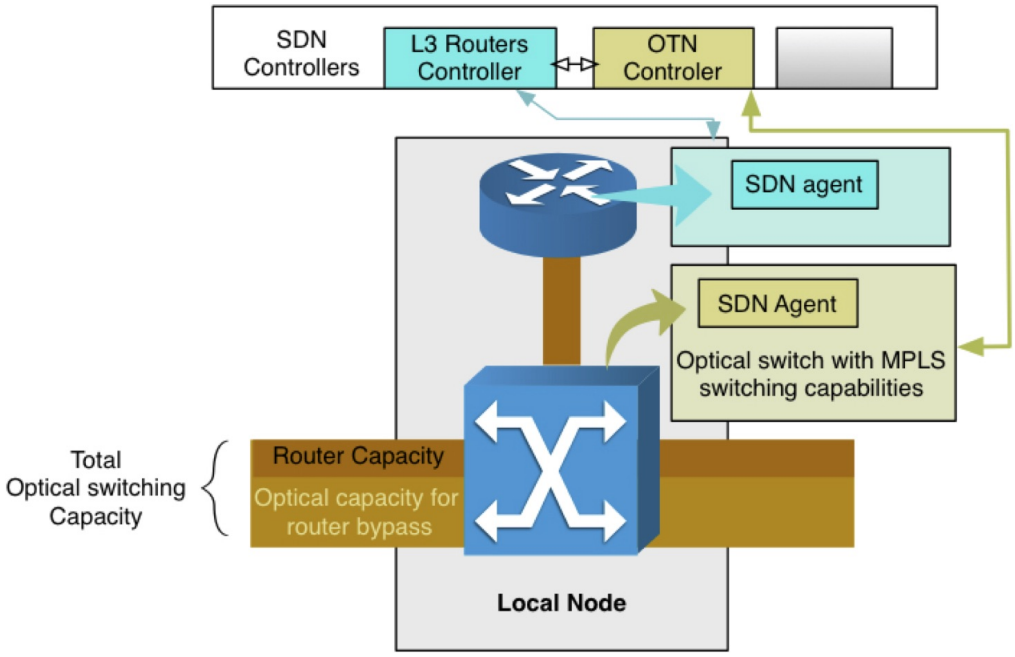


Figure 5.6: Through North-South communication, agents gather information and pass it to controllers. Also, controllers will be able to reconfigure the network by sending commands to agents. East-West messages are exchanged between OTN and routers controllers.

Traditional router bypass underutilizes the network by partitioning links for bypassing traffic. Determining the optimum size of the partitioned channel is a difficult task especially considering the volatile behaviour of Internet traffic. With SDN, partitioned channels will be requested by the application layer. Assuming the application request will be based on maximizing link utilization, bypassing channels will gain the savings from bypassing without compromising link utilization. Figure 5.5 shows a comparison between the traditional bypass and router bypass using SDN. Figure 5.5 (a) shows that the traffic is going through the link without a bypassing channel; bypassing traffic is illustrated in green. Figure 5.5 (a) shows a traditional bypass: the link partitioned on a fixed channel size based on a given value (i.e., average volume of bypassing traffic). Figure 5.5 (b) indicates router bypassing using SDN, where bypassing channels are created based on the source request. Figure 5.5 (a) shows tra-

ditional bypass channels were underutilized because of the high variation of traffic volume during the day.

Current busy-hour Internet traffic is significantly higher than average traffic volume. That volume difference is expected to increase in the future. According to a study [90], busy-hour Internet traffic is growing more rapidly than average Internet traffic. Busy-hour (or the busiest 60-minute period in a day) Internet traffic increased by 51% in 2015, compared with the 29% growth in average traffic. By 2020, the busy-hour will be about five times the average volume [91]. Therefore, assigning constant channel size based on the average use will not be the optimum implementation for router bypassing. Unlike traditional bypassing, applications (traffic sources) will be able to request the router-bypass service from the SDN framework. The network is agile enough to establish a channel as per the client's request using network dynamic orchestration. The clients (i.e., application) will be able to request the service based on parameters which include the time of the service. Figure 5.5 (b) shows how bypassing channels are provisioned for major bypassing sources. Other bypassing traffic will be processed normally by routers.

5.4.2 Capacity Expansion at the Core Node

Given that optical switching offers twice as much capacity at a much lower cost, this paper suggests that switching more traffic by optical switches instead of IP routers can expand the core network capacity [92] [93]. The traffic-aware transport layer will be more efficient in bypassing routers than in previous bypass attempts. Bandwidth can be provisioned for traffic which will be switched through optical equipment without interrupting IP routers. Figure 5.6 illustrates the interaction between global controllers residing in a central device with agents residing on a local node. In SDN, every device in the network has an agent controlled by a global controller. A separate controller will manage the optical layer. Although optical switches have a higher capacity, we propose that the same labelled traffic initiated by an application layer will be switched by an optical switch directly.

5.4.3 Capacity Expansion of the Network

It is well-known that the optical layer has higher capacity and yet the maximum capacity of a traditional IP network is limited to the capacity of routers. Employing router bypass as an SDN service will further utilize the optical network. To calculate the capacity limits at any node, the upper limit is set to the optical capacity and the lower limit is set to the capacity used by routers. That means with no bypassing, the network will operate as a traditional IP/WDM network. With bypassing routers using SDN, the core network will unlock the potential of the optical layer by adding even more switching capacity in the core. The total network capacity can be described as:

$$C = \sum C_{router} + \sum C_{bypass} \tag{5.1}$$

Therefore, the maximum optical bypass capacity can be added to the network capacity (i.e., router's capacity) is

$$C_{bypass} = C_{optical} - C_{router} \tag{5.2}$$

where C is the capacity of the network, C_{router} is the maximum capacity of the core router, $C_{optical}$ is the maximum capacity of optical switches and $C_{O_{bypass}}$ is the available bypass capacity. Assuming always $C_{optical} > C_{router}$, the difference in total optical capacity and the capacity used by routers will be added as the optical capacity available for router-bypass.

The network capacity C will be in a range described as:

$$C_{router} < C < C_{optical} \tag{5.3}$$

This capacity expansion can be extended to the edge nodes when they have a similar optical-MPLS switching capability. The edge router does not usually have such features. Figure 5.7 illustrates the network capacity at the edge and core nodes.

The network capacity in general will be limited by two elements: (a) the edge router capacity, and (b) the optical switching capacity at the core nodes.

Figure 5.7 illustrates capacity gained by using optical router bypassing, In a traditional network, traffic A (in Figure 5.7 (a)) cannot be fully processed in the core router because it is shared with other edge routers traffic B and has limited capacity. However, with a router bypass using SDN (in Figure 5.7 (b)), traffic A is switched optically and the core router is off-loaded from processing that traffic. The overall network will gain extra capacity in the core.

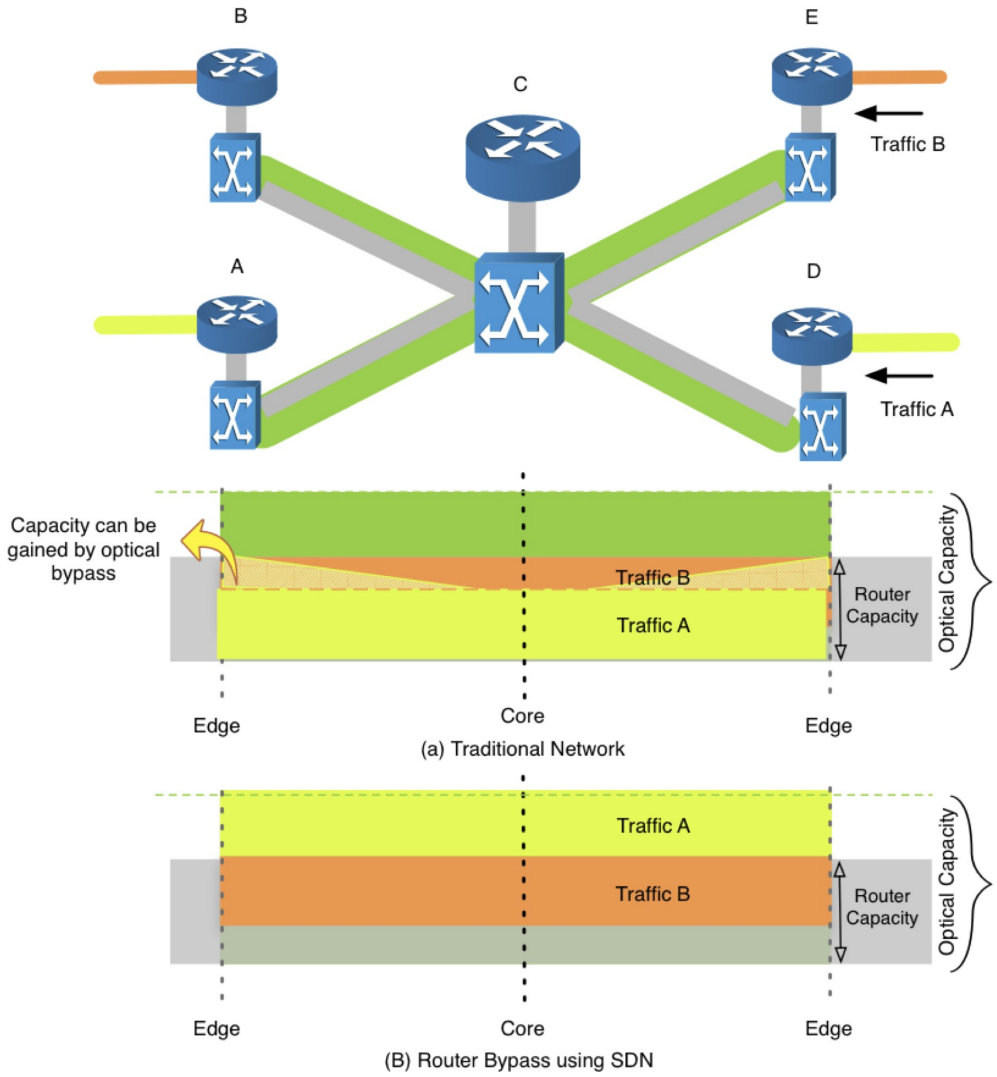


Figure 5.7: Capacity expansion using optical bypass with SDN. (a) Traditional network without bypass is limited by router capacity. Traffic B will be shaped and use what is left from router capacity and the rest will be queued. (b) Optical bypass allows network capacity to expand using optical. Both Traffic A and B will be able to be transferred without queuing by bypass traffic A through optical-bypass channel.

5.5 Simulation and results

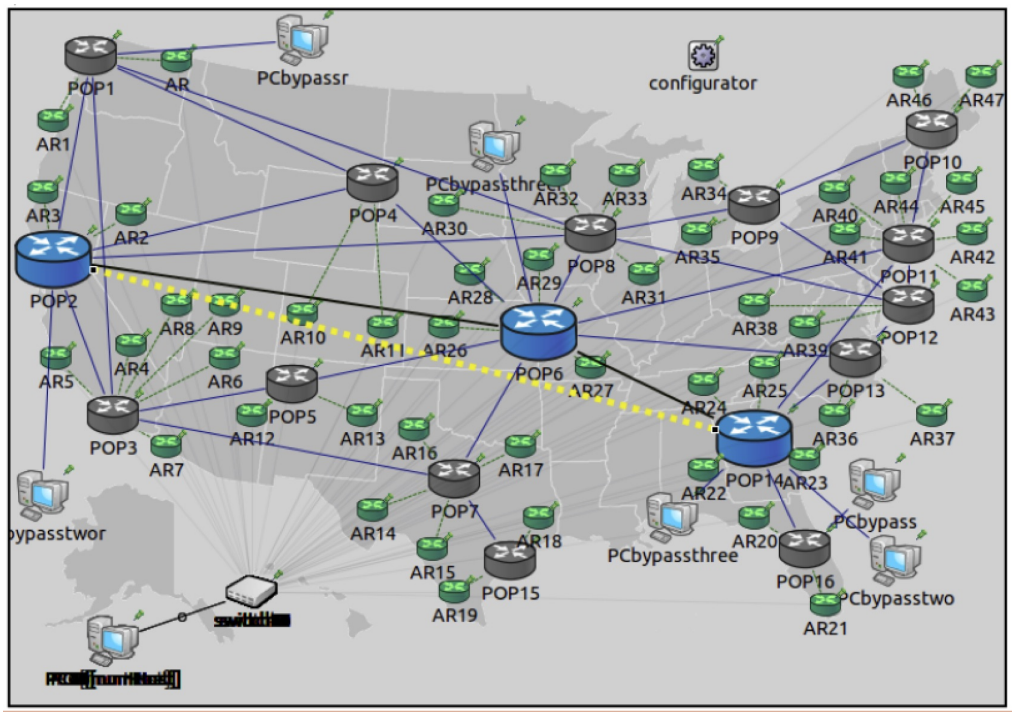


Figure 5.8: Simulation of a core network: bypassing is emulated between blue core routers

In this simulation, we emulate the behaviour of SDN-based bypass traffic behaviour without actually implementing an SDN controlling mechanism. The goal is to compare the bypassing performance of both scenarios: SDN-based and traditional bypass.

To examine the performance of both bypass schemes, a core network simulation has been built to observe and collect statistics using an OMNET++ simulator [60] combined with an INET framework [61]. The simulation consists of 16 core routers (POPs) that are linked together by 40 Gbps channels. Core routers are connected to 47 access routers ARs as shown in Figure 5.8. Every AR is connected to a switch serving 10 clients as traffic generators. Every client sends traffic uniformly to other clients in the network at a low rate. The bypassing links through the transport network are simulated as a direct link for simplicity, with modification of the delay factor.

To form the bypass traffic shown in Figure 5.9, traffic generators are linked to

POP14 to provide the bypass traffic targeting POP2. That traffic will bypass POP6 by partitioning the link between POP14 and POP6 as well as POP6 and POP2. In Figure 5.9, the yellow line represents the channel size of a traditional bypass while the orange line represents the channel size of the proposed SDN-based bypass. The bypass link will directly connect POP14 to POP2 with an OTN connection as illustrated in Figure 5.8 (yellow link). Time values are normalized to T= 10^{-6} s. The simulation has been run five times with different initial seeds and the average values were calculated.

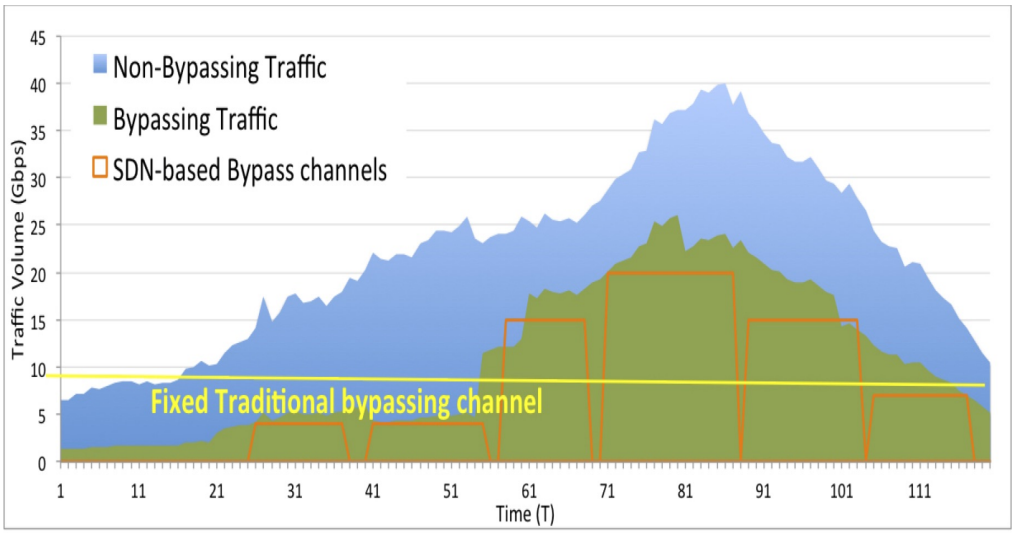


Figure 5.9: Provisioned bypass channels: traditional vs. SDN-based bypass

There are multiple models that emulate Internet traffic behaviour over period of time. In our bypass approach, we assumed that traffic will be bypassed after being aggregated. Therefore, less variation (volatility) will occur in the gathered traffic. However, if the bypassing channels are established based on individual resources then higher inefficiencies will be expected. These inefficiencies will depends on which traffic models is being used and how closely the equipment will be able to adapt to that model. Adapting real time traffic is virtually impossible which is why we assumed traffic will be aggregated. We used the Google daily traffic usage pattern as a practical guideline for the generated traffic with 15% standard deviation to add randomness. Considering different traffic models for router bypass will be an interesting topic for further investigation. In our simulation, we have assumed that the OTN layer will optimize and manage the bandwidth resources. The upper limit of optimization

will be limited by the OTN layer.

Figure 5.9 mimics the traffic pattern for a day in a shorter time-frame [69] going from POP14 to POP6. The traffic consists of two parts: (a) bypass or transit traffic (in green) going to POP2, and (b) traffic processed by POP6 locally (non-bypass). The overall traffic theme imitates the normalized value of traffic demand reported for Google in Canada during one day [69]. In the same figure, both bypass schemes are projected on top of bypass traffic. We are assuming that SDN channels are initiated based on application requests.

Figure 5.10 and 5.11 show the two schemes of provisioned channels (a)

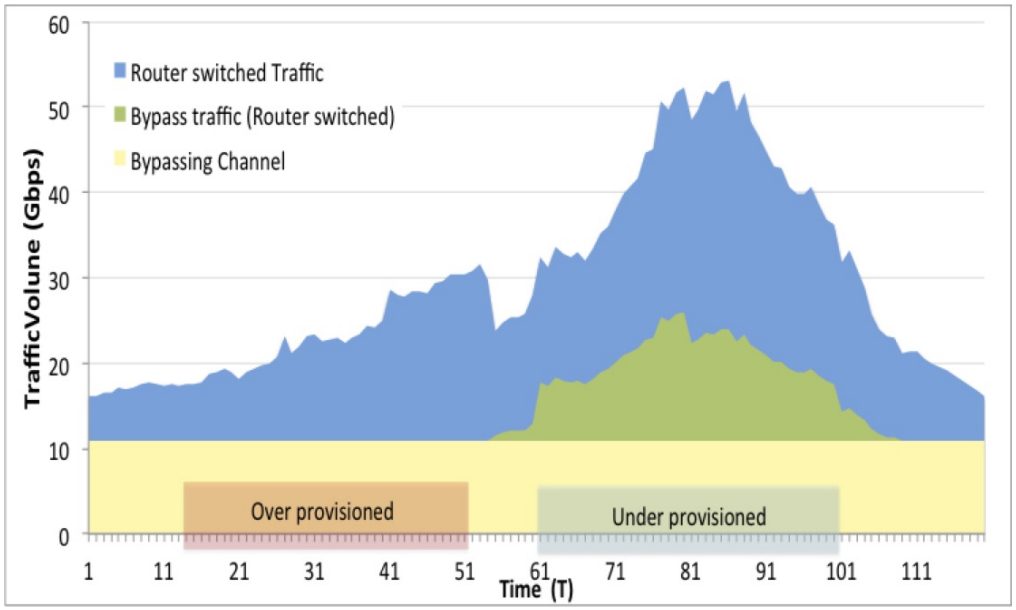


Figure 5.10: Bypassing channel provisioning: (a) Fixed traditional provisioning

traditional bypassing, and (b) SDN-based bypassing. A traditional bypassing channel is fixed size based on average volume. The problem with static provisioning is that it does not track the change in traffic volume over time which means that most of the time it is not efficient or optimized.

The percentage of transit traffic to the total volume changes over time. The average of transit traffic over time has been calculated as 11 Gbps. Traditional bypassing creates a fixed-size bypassing channel based on calculations (i.e., average volume of transit traffic) by partitioning the main link; SDN creates bypassing channels based on client requests. In the simulation, we assumed

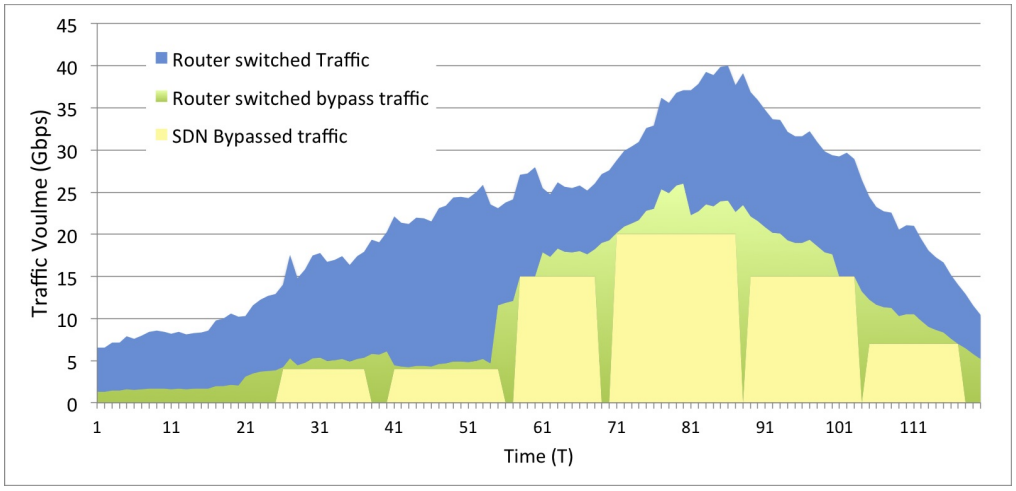


Figure 5.11: (b) Adaptive SDN-based provisioning

T	Transit	SDN b	Trad. b	local traffic	SDN local cap.	Trad. local cap.
28T	4.1	5 Gbps	11 Gbps	11 Gbps	35 Gbps	29 Gbps
30T	4.2	5 Gbps	11 Gbps	23 Gbps	35 Gbps	29 Gbps
45T	3.6	4 Gbps	11 Gbps	31 Gbps	36 Gbps	29 Gbps
50T	3.3	4 Gbps	11 Gbps	29 Gbps	36 Gbps	29 Gbps
60T	18	15 Gbps	11 Gbps	11 Gbps	25 Gbps	29 Gbps
80T	20	20 Gbps	11 Gbps	14 Gbps	20 Gbps	29 Gbps
95T	17	15 Gbps	11 Gbps	12 Gbps	25 Gbps	29 Gbps
111T	8	6 Gbps	11 Gbps	4 Gbps	34 Gbps	29 Gbps

Table 5.1: Traffic samples with traditional router bypass and SDN bypass. SDN bypass channels are adaptive to transit traffic. Transit traffic is sometimes higher than the SDN bypass channel capacity to indicate that some transit traffic doesn't justify expanding bypass channels further.

most of the transit traffic is generated based on SDN requests and so bypass channels are sized based on these requests.

5.5.1 Traditional bypass over-provisioning vs. SDN-based bypass

Figure 5.10 highlights (in red) where the channel is over-provisioned and where bypass traffic volume is less than the bypassing channel. At the under-provisioned area (light blue), the potentially higher savings can be achieved by assigning more capacity to the bypass channel. We will compare the performance of both areas with SDN-based bypassing. Figure 5.10 shows the

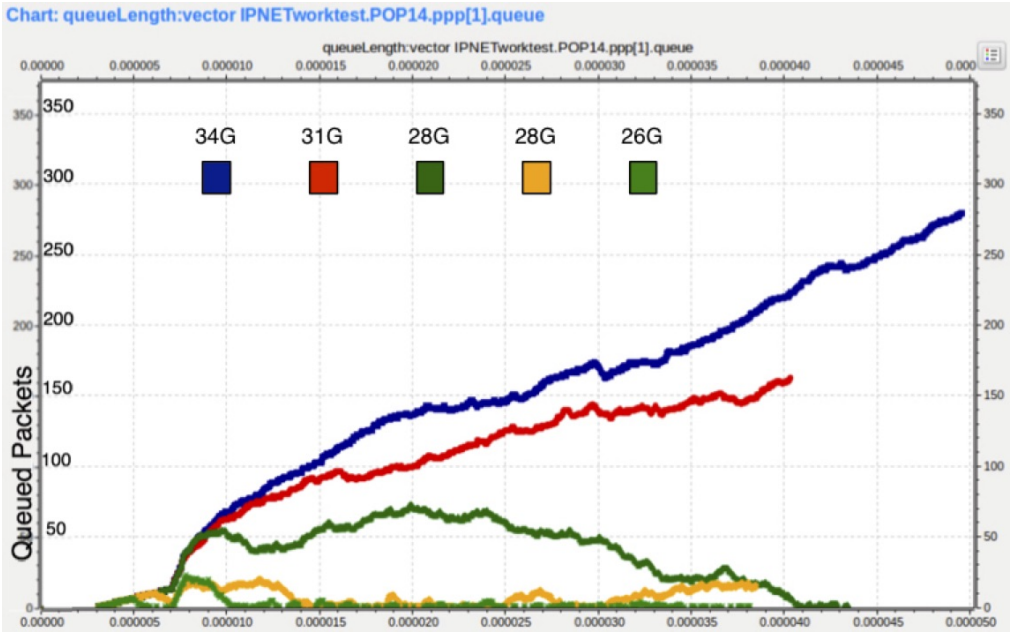


Figure 5.12: Queuing length with traditional bypassing

SDN bypassing scheme where major bypassing clients requested a bypassing service. The underlying SDN network provisions the bypass channels to accommodate the requests. In contrast to traditional bypassing, the SDN bypass channel capacity will track the volume demand. Therefore, it should be neither over-provisioned nor under-provisioned with SDN. An assumption in the SDN model is that the service will offer channel sizes to make sure the clients fully utilize the bypassing channel.

Table 5.1 shows the values used in the simulation to emulate the over-provisioning area. We note that the SDN framework follows intelligent bandwidth allocation for traffic unlike the traditional bypass. Even when local traffic volume is high, the traditional bypass does not change the bypass channel size (11 Gbps) to allocate enough bandwidth to local traffic even if it is not needed by the transit traffic. The size of all packets is set to 100 bytes. UDP is the type of generated traffic with various port numbers (6000, 9000 and 1000).

When the traditional bypass channel is in the over-provisioning area (red), the bypass channel is given more capacity than needed and this might cause congestion in the link when local (non-transit) traffic volume is over the re-

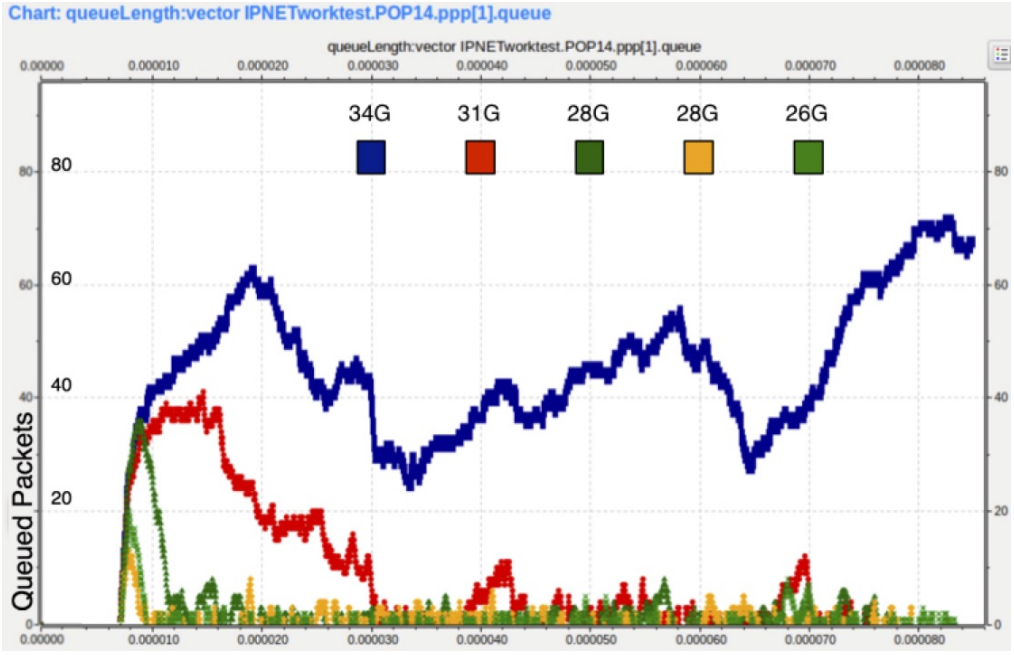


Figure 5.13: Queuing length with SDN-based bypassing

maining capacity. The congestion leads to higher traffic delays and weakens the link performance in general. Figure 5.14 shows a comparison between traditional bypassing and SDN-based bypassing in the over-provisioning scenario. Queuing length was measured at POP14 (at the link between POP14 and POP6) at multiple points of time in the red area (various local traffic volumes) as shown in Figure 5.14 and Figure 5.15. With SDN, the bypassing capacity offered is based on source needs which are about 5 Gbps; traditional bypassing follows the constant average 11 Gbps. Because local traffic was higher than bypass traffic, traditional bypassing performed worse than SDN. With the SDN bypassing scheme, the maximum queuing length is around 73 packets. In contrast, in the traditional bypassing, it reached the 280-packet mark and kept rising dramatically because of the link congestion. Packets are dropped if queue capacity exceeds buffer size. The buffer size in our experiment is set to 500 packets. In traditional bypassing, congestion is severe at both local traffic volume of 35 Gbps and 31 Gbps; congestion has been avoided in SDN. Assigned capacities in traditional bypassing are not optimized when traffic behaviour diverges from the mean.

Figure 5.16 shows the mean delay on the link at various traffic volumes. In

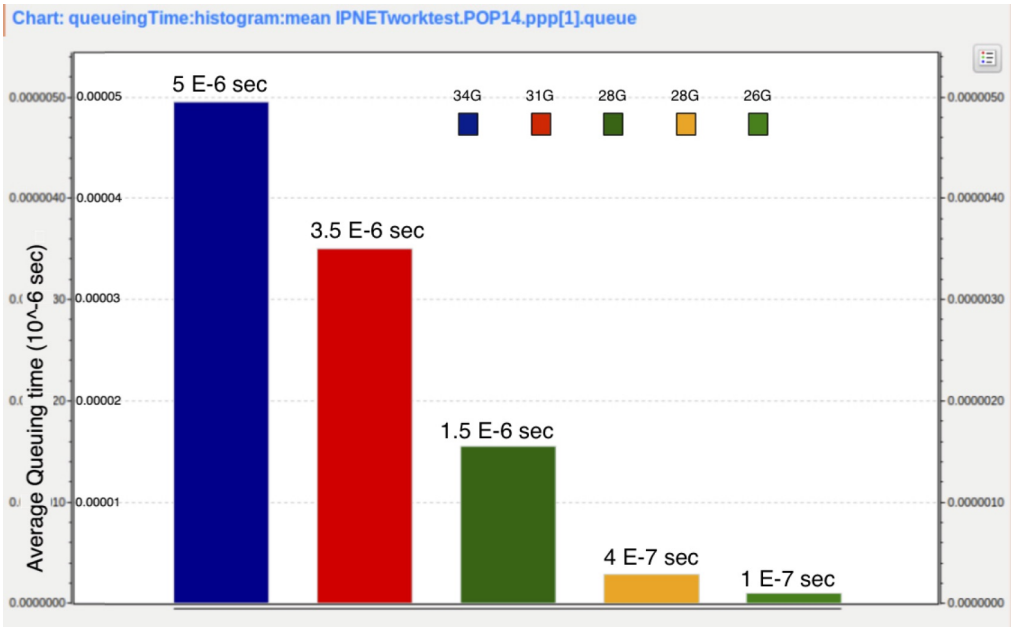


Figure 5.14: Average queuing time with traditional bypassing

traditional bypass, as shown in Figure 5.16, the delay was significantly higher than SDN-based bypass at the same traffic volume as shown in Figure 5.17. At high traffic volumes, SDN delay was less. At lower volumes, SDN bypassing was better by a factor of 10 because the congestion had been avoided.

5.5.2 Overall bypassing performance

Comparing both schemes in the over-provisioning area, traditional bypass under-provisioned channels, when traffic volume exceeds clearly the fixed channel size. This means that the traditional bypass in that area will be fully utilizing the links. However, traditional bypass will not be able to bypass more traffic which should be bypassed to maximize savings. An optimum bypassing scheme should fully utilize the bypassing links and bypass as much transit traffic as possible. If both schemes can reach a similar level of utilization, then the scheme with the higher bypassed traffic volume will be optimual. SDN bypassing channels will be able to accommodate more traffic by provisioning larger channels when it is needed.

Figure 5.16 illustrates cumulative bypassed traffic volume over time for both schemes. For almost half the time of the experiment, the traditional scheme

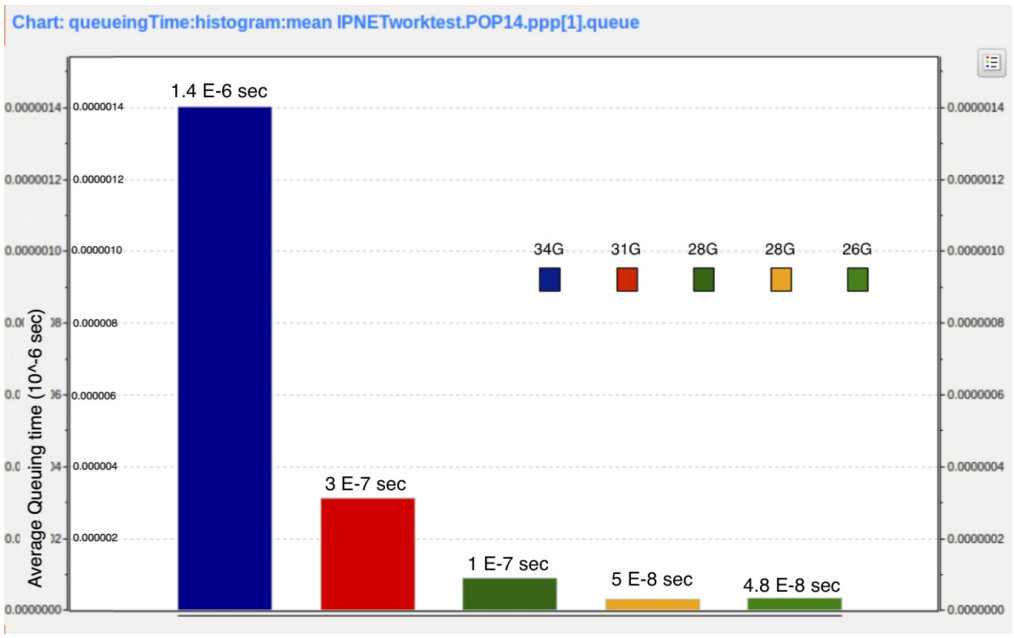


Figure 5.15: Average queuing time with SDN proposed bypassing

bypassed more traffic; however, in the long run, the SDN scheme bypassed more traffic. The traditional scheme bypasses more traffic in the beginning by compromising link utilization, as shown in the previous section. The SDN scheme was able to outperform the traditional scheme by provisioning larger channels where needed. With SDN, the network operator will be able to maintain high link efficiency and yet capture a large volume of bypass traffic. To measure the efficiency of both bypassing schemes, we describe the dimensionless bypassing efficiency factor as:

$$\gamma = \frac{\text{Volume of bypassed traffic in bps}}{\text{Bypass channel size in bps}} \tag{5.4}$$

Figure 5.17 shows the efficiency factor for in both bypass schemes. With SDN, high efficiency (up to 99%) can be reached by tracking bypass traffic volume. That efficiency requires dynamic provisioning tools with minimum overheads. An assumption made in the experiment is that when the SDN bypass channel is provisioned, it will be fully utilized.

In addition, with SDN, the utilization factor can be monitored and controlled by the network operator. Traditional bypassing had a lower bypass efficiency

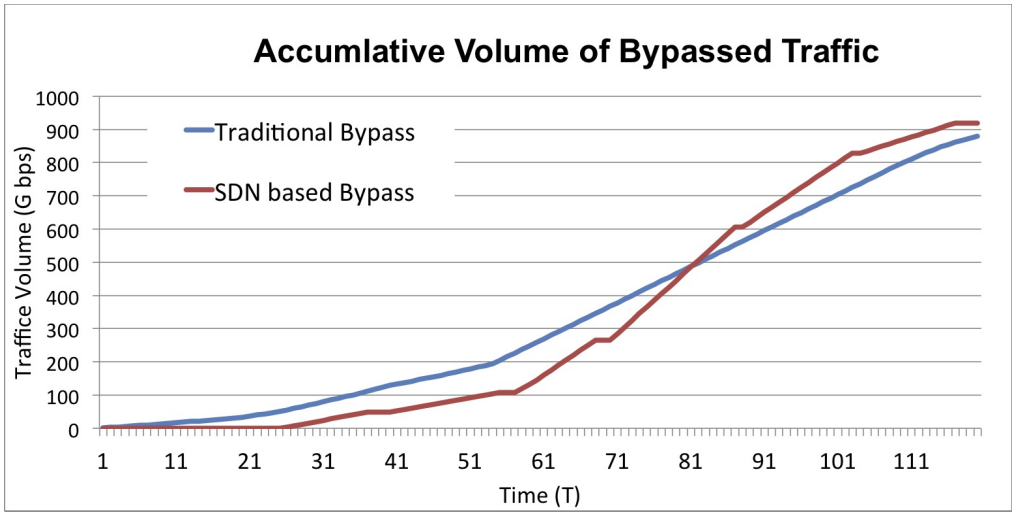


Figure 5.16: Cumulative volume of bypassed traffic in both bypass schemes

because the provisioned bandwidth does not track highly volatile traffic and most of the time ends up with over-provisioned channels.

5.5.3 Expanding Node Capacity

Lowering link utilization by router bypass is an undesirable side effect. To maintain link utilization around 60% (on average) in our experiment, the bypass link cannot be expanded beyond 11 Gbps. However, with SDN, bypass channels will be established on demand; therefore, the bypass should maintain a high level of utilization. When router capacity is fully utilized during rush-hours, bypass traffic can utilize the remaining optical capacity. With 80 Gbps of optical capacity, Figure 5.18 shows how the node capacity can be expanded with an SDN-based bypass by up to 56% compared with the traditional bypass, without compromising link usage. The granularity offered in OTN allows that to be even more feasible with lower bypass-traffic rates.

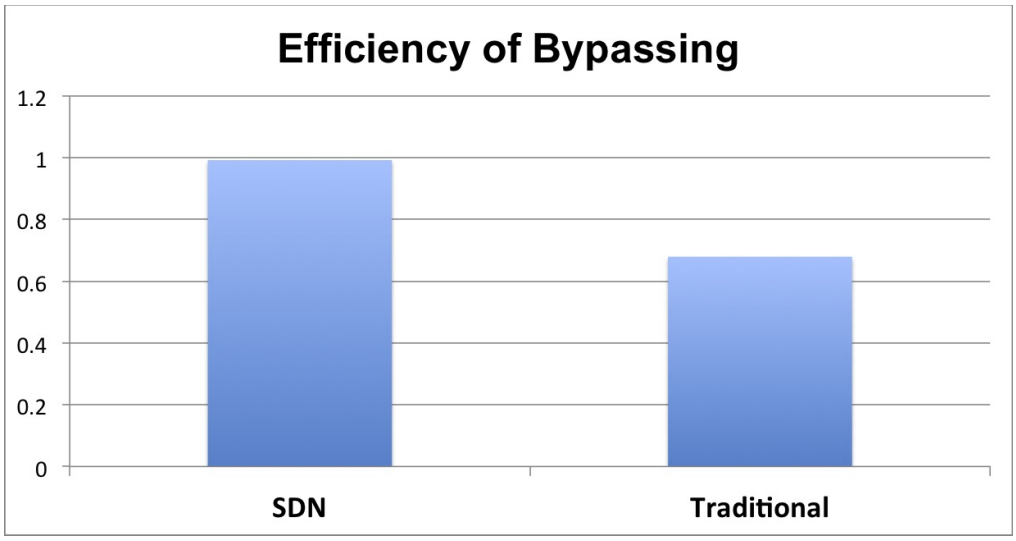


Figure 5.17: Average efficiency of both bypassing schemes

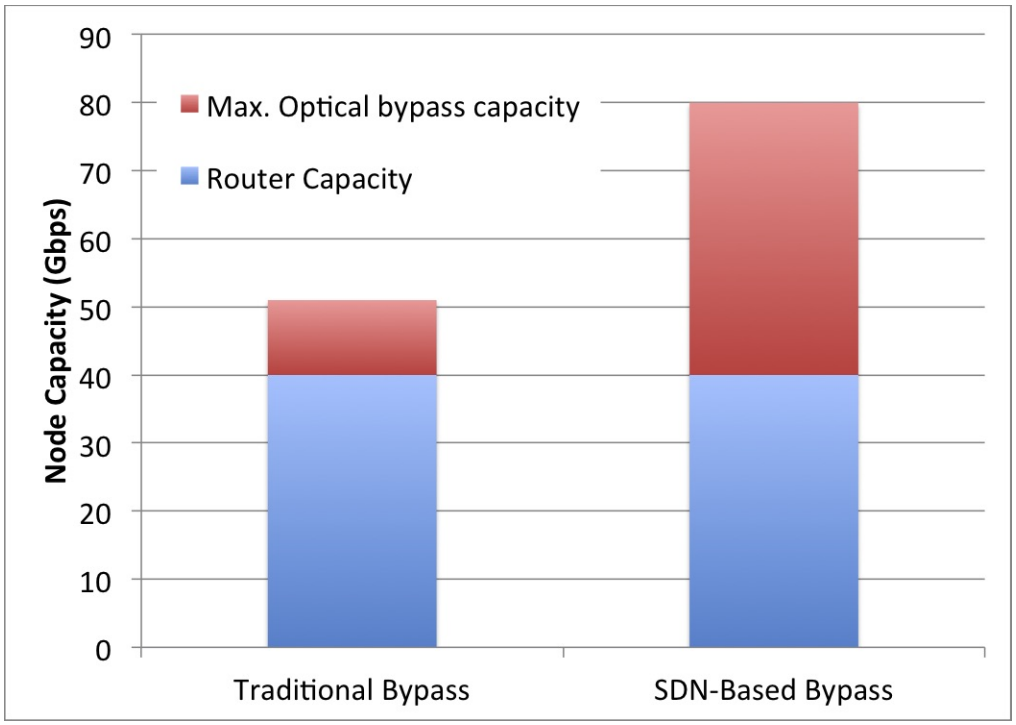


Figure 5.18: Maximum node capacity for both bypass schemes

5.6 Summary

OPEX and CAPEX savings made by using router bypass was shown in many studies. However, volatile traffic behaviour during the day made the concept of fixed-channel bypassing less efficient. Traditional bypassing has not been widely adopted because of its lack of agility in a dynamic environment. With the introduction of SDN on top of the advances in the optical transport technologies, this study has explored the concept of router bypassing based on an SDN framework.

By using SDN-based bypassing and OTN granularity, this study showed that the network can improve in three aspects: (a) the network capacity can be expanded beyond the router capacity, (b) the bypassing process will increase savings with a higher efficiency (up to 99%), and (c) more visibility occurs in the bypassing process. Instead of static and fixed bypassed channels, the bypassing service is requested by applications and the network will offer the service where it is more economical. A simulation experiment was developed using OMNET++ and the INET framework. The results showed that SDN bypassing is more efficient and can increase bypassing savings without compromising the link utilization. The experiment illustrated that, overall, SDN bypassing is more efficient than traditional bypassing. With traffic model, we assumed it is aggregated and distributed over a day, it would be useful to test other traffic models. We see challenges in adaptation where proprietary implementations must give way to a common and open framework.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

By working on enhancements to the router bypassing concept, we provided improvements to the core network efficiency, which will advance the evolution of the Internet. In addition, improving network efficiency will reduce network costs and the carbon footprint by saving power consumed in the switching process. We predict that enhancing the router bypassing mechanism will lead to integrated adaptive electronic/optical switching networks in the future.

In this dissertation, we have investigated the router bypass concept with new enhancements in the optical transport layer and further introduced the impact of SDN and automated provisioning on the bypass.

In Chapter 2, we studied the various technologies that have been used to deploy router bypass as a means to strengthen the layout and foundation of our proposals. We explained the foundations of network principles and how data moves across networks. We explored technologies WDM: DWDM CDWM , the architecture of IP over WDM networks, and the advancements in ROADMs which play an important role in future network automation.Then, we explained how MPLS as a technology can be run over the transport layer in sub-lambda switching concept; tagging channels instead of packets makes traffic switching more efficient and scalable.

We studied the advancements in the Optical Transport Layer. These allow transport, multiple, switch, manage, supervise optical channels carrying client signals. Many features such as Hitless Adjustment OTN has been explored, where the transport network can adjust the channel capacity without tearing it down. ODUflex is another adaptive channel speed which can be used even for transporting packets as long as it is within the allowed speed. OTN offers a wide range of channel capacities to accommodate a variety of traffic types and speeds which have been incorporated in our router bypass solutions. Network automation technologies such as the Software Defined Networking (SDN) and Network Virtualized Function (NVF) are explored and incorporated in the router bypass context in Chapter 5.

In Chapter 3, router bypass as a concept is investigated in different ways of deployment: over WDM and over OTN, explaining the advantages and savings to be realized. Potential savings are highlighted with wide deployments of router bypass. The issue of losing statistical multiplexing as a bypass drawback is investigated which is caused by portioning the transport channel. Our guidelines to enhance the router bypass mechanism are highlighted such as: granular bypassing and traffic-based adaptive bypass. A preliminary simulation is built with results being collected. We described the main INET submodule (blocks) of our simulation which are used to emulate routers and channels.

In Chapter 4, motivated by the enhancements in the transport layer, we proposed the Adaptive Router Bypass using Feedback Adjusted OTN, and Enhanced Router Bypass using Fine Granularity. In these proposals, we laid out how the router bypass can be adaptive to network traffic by measuring traffic volume and adjust the bypass channel according to that traffic volume. We studied the impact of channel granularity on the bypass performance, with finer granularity channels showing superior results but at the cost of channel re-adjustment. However, the equipment adjustment speed might be the limitation factor.

Since the network usage varies during the day, we proposed a dynamic granu-

larity. In a dynamic granularity bypass, the network will use past information (i.e., traffic burstiness) to anticipate the appropriate bypass channel capacity. Additional capacity can be added (or reduced) based on a calculated slope of past traffic volume. Then simulations and analysis showed that dynamic granularity could maximize bypass savings with a lower impact on throughput (which is typical impacted negatively by traditional router bypass).

In Chapter 5, traffic provisioning will be established and controlled by traffic (i.e., applications) unlike the assumption in the previous chapter. This concept is offered in SDN where applications will request network services and the network will accommodate these requests based on availability. With this paradigm change, we investigated the impact of SDN on router bypass. Assuming that applications will request channel capacity in an efficient fashion, we showed that router bypass with (SDN) can achieve even better results without the need of complex traffic prediction algorithms. Offering router bypass as an (SDN) service will integrate router bypass to the network in demand-based fashion. Unlike traditional bypass, service providers will able to pass the cost benefits of bypass to customers in a more transparent method of network provisioning.

6.2 Future Work

For the future work planned beyond this dissertation, various open issues of importance still exist:
1. Cross-multilayer switching: the integration between the optical transport and IP layer is important for router bypass enhancement. For instance, in SDN, most deployments are done by using multiple controllers: one for the optical and one for the IP layer. The shared information between these controllers (or layers) will help to develop a more consolidated integrated core network. The integration will enhance the provisioning efficiency. Having multiple controllers on separate physical equipment represents a challenge for the industry in building an integrated network. Open source-based controllers

with protocols for the controllers to share information will allow equipments to provision the network within three layers in a consolidated fashion.
Also required is a comprehensive protocol to determine when provisioned bandwidth should be changed in the optical layer or IP layer. IP addressing should be an independent element from the switching aspect and stay stable more often as long there is a connectivity to recorded routes. Then, switching part can be done through optical or typical packet switching. This area requires a more comprehensive protocol to be built.
2. Building software to implement the proposed protocols. Whether feedback-based bypass or offer a bypass as an SDN service, these solutions need to be programmed and run over open source enabled equipment.
3. Building hardware to embed the router bypass techniques proposed in this dissertation. The building of a feedback based router bypass requires gathering data about traffic volume on a periodic basis. Then, the adjusting of the optical channels will occur based on the parameters previously mentioned in this proposal. Adjusting channels on a regular basis can be challenging especially if the equipment requires a long time to converge. Building a hardware experiment will allow this proposal to come one step closer to being implemented on a commercial scale.
4. Studying the correlation between traffic type and router bypass that correlation is worth further research. Data traffic and flows have multiple types and characteristics. These characteristics vary in terms of time sensitivity, the amount of required acknowledgment and whether packets are large or small. Observing the impact of router bypass on various type of traffic will allow informed decision to be made in terms of designing router bypass algorithms.

6.3 Contributions

The main contributions of this dissertation are:
- Reviewing the efficiency issues in current core networks
- Reviewing core network technologies.
- Illustrating benefits and drawbacks of traditional router bypass
- Studying OTN and how it can be used to enhance router bypass.
- Proposing multiple adaptive techniques to enhance router bypass:

  - Using Granularity factor in OTN to enhance router bypass
  - Traffic-based adaptive router bypass: based on changes in traffic volume
  - SDN based router bypass

- Building up flow charts for adaptive router bypass for every technique
- Emulation multiple router bypass scenarios based OMNET++ and INET environment.
- Presenting the enhancement and the trade-off in the proposed router bypass techniques.

6.3.1 Publications

- F. A. Ghonaim, T. E. Darcie and S. Ganti, "Adaptive router bypass using feedback adjusted OTN", 2015 22nd International Conference on Telecommunications (ICT), Sydney, NSW, 2015, pp. 123-127. doi: 10.1109/ICT.2015.7124669
- F. A. Ghonaim, T. E. Darcie and S. Ganti, "Enhanced router bypass using fine granularity transport channels", 2015 International Conference on Computer, Information and Telecommunication Systems (CITS), Gijon, 2015, pp. 1-5. doi: 10.1109/CITS.2015.7297753
- (Accepted) F. A. Ghonaim, T. E. Darcie and S. Ganti, "Impact of SDN on Router Bypass" Journal of Optical Communications and Networking March 2018

Bibliography

[1] Cisco Visual Networking Index (VNI). "The Zettabyte Era - Trends and Analysis." Internet: www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.html, Jun. 10, 2014 [Feb. 15, 2015].

[2] Cisco Visual Networking Index. "Forecast and Methodology, 2014-2019 White Paper." Internet: www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html May 27,2015 [Oct. 21, 2015].

[3] S. Stroh, G. Schroder and F. Grone. "Keeping the Data Center Competitive Six Levers for Boosting Performance, Reducing Costs, and Preparing for an On-Demand World" Available: www.strategyand.pwc.com/media/file/Keeping_Data_Center_Competitive.pdf 2009. [Oct. 20, 2015].

[4] P. Van Mieghem. (2006, March 9) *Performance analysis of communications networks and systems.* (1^{st} edition). Cambridge, 2006, pp. 357-358.

[5] Innovation Observatory Ltd. "Router off-load strategies. When is router off-load an attractive option." *Report* (October, 2011), pp. 4-12.

[6] D. Webster. "Prime Time for Broadband: Part 2: Announcing the 2010 Cisco VNI Usage Study" Internet: http://blogs.cisco.com/sp/prime-time-for-broadband-part-2-announcing-the-2010-cisco-vni-usage-study Oct. 27, 2010 [Feb. 15, 2015].

[7] P. Belotti *et al.* "Transport networks at a crossroads," Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference , vol., no., pp.1,3, 6-10 March 2011.

[8] S. Gangxian and R.Tucker. "Energy-Minimized Design for IP Over WDM Networks," *Optical Communications and Networking, IEEE/OSA Journal*, vol.1, no.1, pp.176,186, June 2009.

[9] Nokia Siemens Network. "Optical Transport Network Switching:Creating efficient and cost-effective optical transport networks" Internet: `http://networks.nokia.com/system/files/document/optical_transport_network_switching.pdf` [Feb. 15, 2015].

[10] J. Baliga, K. Hinton, R.S. Tucker., *Energy Consumption of the Internet* , Optical Internet, 2007 and presented at the 2007 32nd Australian Conference on Optical Fiber Technology. COIN-ACOFT 2007. Joint International Conference (June 2007) 1-3.

[11] Bone *BONE project (2009). WP 21 topical project green optical networks* (Report on year 1 and updated plan for activities). NoE, FP7-ICT-2007-1 216863 BONE project, Dec. 2009.

[12] http://newsroom.cisco.com/ Cisco Visual Networking Index Study - *New Cisco Study Reveals Peak Internet Traffic Increases Due to Social Networking and Broadband Video Usage* - October 2009

[13] S. Lambert, Ward Van Heddeghem, Willem Vereecken, Bart Lannoo, Didier Colle, and Mario Pickavet, *Worldwide electricity consumption of communication networks* Optics Express 2012 (Vol. 20, Issue 26, pp. B513-B524)

[14] H. K.; Benjamin, S.; Geron, A.; Katz, G.; Stepanov, S.; Margalit, N.; Mesh, M., *"A high end routing platform for core and edge applications based on chip to chip optical interconnect,"* Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC), 2013 , vol., no., pp.1,3, 17-21 March 2013

[15] *Greenhouse Gas Equivalencies Calculator.* please go to http://www.epa.gov/cleanenergy/energy-resources/calculator

[16] D. Mead, *The Next Five Years of Explosive Internet Growth, in Seven Graphs*. please go to http://motherboard.vice.com/blog/the-next-five-years-of-explosive-internet-growth-in-seven-graphs

[17] Cisco CRS-1 16-Slot Single-Shelf System, Product Data sheet

[18] www.stoke.com *Will Small Packets Degrade Your Network Performance?*

[19] Y. Chen, O. Oguntoyinbo., *Power efficient packet classification using cascaded bloom filter and off-the-shelf ternary CAM for WDM network* Computer Communications 32 (2009) 349-356. Worldwide electricity consumption of communication networks

[20] R. Tucker, *Will optical replace electronic packet switching* SPIE Newsroom 2007

[21] Global Internet Phenomena Report, Sandvine 2013

[22] S. Melle, D. Perkins and C. Villamizar, "Network Cost Savings from Router Bypass in IP over WDM Core Networks," Optical Fiber communication/National Fiber Optic Engineers Conference, 2008. OFC/NFOEC 2008. Conference on, San Diego, CA, 2008, pp. 1-10.

[23] S. Das, G. Parulkar and N. McKeown, "Rethinking IP core networks," in IEEE/OSA Journal of Optical Communications and Networking, vol. 5, no. 12, pp. 1431-1442, Dec. 2013. doi: 10.1364/JOCN.5.001431

[24] S. Kasera *Atm Networks Concepts And Protocols* - second edition 2006

[25] S. Jane M.; Saleh, A.A.M., "The value of optical bypass in reducing router size in gigabit networks," Communications, 1999. ICC '99. 1999 IEEE International Conference on , vol.1, no., pp.591,596 vol.1, 1999

[26] R. L. Freema *Telecommunication System Engineering* - fourth edition 2004

[27] *Average Web Page Size Triples Since 2008* please go to http://www.websiteoptimization.com/

[28] A. G. Blank *TCP/IP JumpStart: Internet Protocol Basics* - Second edition 2002

[29] W. Odom *Computer Networking First-Step* - 2004

[30] D. Medhi, *Network Routing: Algorithms, Protocols, and Architectures*

[31] A. Perez, *IP, Ethernet and MPLS Networks: Resource and Fault Management* - 2011

[32] Z. Wang et al. *Vector Address Switching: An Energy-Effective Next-Generation Switching Technology for Internet*

[33] V. Alwayn *Optical Network Design and Implementation* - 2004

[34] `http://lightriver.com/transport-connectivity/roadm-dwdm` [April. 06, 2018].

[35] K. H. Liu *IP over WDM*, John Wiley and sons, 2003, p. 155-173

[36] *http://www.itu.int/ITU-T/2001-2004/com15/otn/definitions.html*

[37] M. Raman et al. *Next Generation Transport Networks - Data, Management and Control Planes* 2005

[38] *Optical Transport Network (OTN) Tutorial*. Available at: `https://www.itu.int/ITU-T/studygroups/com15/otn/OTNtutorial.pdf`

[39] ITU-T G.709/Y.1331, *Interfaces for the Optical Transport Network (OTN)* 2003

[40] TPack. Odu0 and oduflex *a future-proof solution for OTN client mapping* White paper, 2010.

[41] Rec. ITU-T G.7044/Y.1347 (10/2011). *"Hitless adjustment of ODUflex(GFP)"*

[42] Vern Paxson and Sally Floyd, *"Why We Don't Know How To Simulate The Internet"* - Network Research Group Lawrence Berkeley National Laboratory - CiteSeerX

[43] http://www.exar.com *"ODUflex in Detail : Transporting Any Client Signal in the OTN"*

[44] ITU-T. "Hitless Adjustment of ODUflex(GFP) (HAO) G.7044/Y.1347" Available: http://www.itu.int/rec/T-REC-G.7044/en Oct. 2011. [Feb. 15, 2015].

[45] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky and S. Uhlig, "Software-Defined Networking: A Comprehensive Survey," in Proceedings of the IEEE, vol. 103, no. 1, pp. 14-76, Jan. 2015.

[46] Open networking foundation *Software-Defined Networking (SDN) Definition* Internet:https://www.opennetworking.org/sdn-resources/sdn-definition [Jan. 10, 2017].

[47] S. Fogarty, *7 Essentials Of Software-Defined Networking* Internet:http://www.networkcomputing.com/cloud-infrastructure/7-essentials-software-defined-networking/1672824201, 11/19/2013 [Jan. 15, 2017].

[48] SDx central, *How Does Software-Defined Networking or SDN Work?* Internet:https://www.sdxcentral.com/sdn/definitions/what-the-definition-of-software-defined-networking-sdn/ [Jan. 15, 2017].

[49] Ciena.com *What is NFV* Internet:http://www.ciena.com/insights/What-is-NFV_prx.html [Jan. 21, 2017].

[50] Rajendra Chayapathi, Syed F. Hassan, Paresh Shah, *Network Functions Virtualization (NFV) with a Touch of SDN* Nov 28, 2016.

[51] Cisco.com *NFV - Network Functions Virtualization* Internet:http://www.cisco.com/c/en/us/solutions/service-provider/network-functions-virtualization-nfv/index.html [Jan. 24, 2017].

[52] E. Zouganeli. *Optical networks: From point-to-point transmission to full networking capabilities* Telektronikk, February 2005.

[53] M. Chamania, M. Caria, A. Jukan *"Achieving IP routing stability with optical bypass"* - Optical Switching and Networking (2010)

[54] S. Das, G. Parulkar, N. McKeown *"Rethinking IP Core Networks"* - IEEE Optical Communications and Networking, Dec 2013

[55] Xtera Communications, Inc. *"http://www.xtera.com/"* news - March 4, 2013

[56] http://www.cisco.com/ *"Cisco CRS 16-Slot Single-Shelf System Data Sheet"*

[57] P. Van Mieghem, *Performance Analysis of Computer Systems and Networks* Cambridge (2006)

[58] R. S Tucker, Rajendran Parthiban, Jayant Baliga, Kerry Hinton, Robert W. A. Ayre, and Wayne V. Sorin, *Evolution of WDM Optical IP Networks: A Cost and Energy Perspective* JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 27, NO. 3, FEBRUARY 1, 2009

[59] http://www.cisco.com/ *"Cisco CRS 4-Port 40GE LAN/OTN Interface Module"*

[60] OMNeT++ Discrete Event Simulator. Available at: `omnetpp.org` [March. 15, 2015].

[61] INET Framework - OMNeT++ Available at: `inet.omnetpp.org`

[62] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modelling," in IEEE/ACM Transactions on Networking, vol. 3, no. 3, pp. 226-244, Jun 1995.

[63] G. Mansfield, T. K. Roy and N. Shiratori, "Self-similar and fractal nature of Internet traffic data," Proceedings 15th International Conference on Information Networking, Beppu City, Oita, 2001, pp. 227-231.

[64] Zhang, Yin et al. ?On the characteristics and origins of internet flow rates.? SIGCOMM (2002).

[65] Cisco systems Inc. "Best Practices in Core Network Capacity Planning: Architectural Principles of the MATE Portfolio of Products". Internet: `http://www.cisco.com/c/en/us/solutions/collateral/service-provider/quantum/white_paper_c11-728551.pdf` 2013 [March 12, 2015].

[66] F. A. Ghonaim, T. E. Darcie and S. Ganti, *Adaptive router bypass using feedback adjusted OTN* 2015 22nd International Conference on Telecommunications (ICT), Sydney, NSW, 2015, pp. 123-127. doi: 10.1109/ICT.2015.7124669

[67] N. Spring, R. Mahajan, and D. Wetherall, ?Measuring ISP topologies with rocketfuel,? in SIGCOMM, Aug. 2002.

[68] F. A. Ghonaim, T. E. Darcie and S. Ganti, "Enhanced router bypass using fine granularity transport channels," 2015 International Conference on Computer, Information and Telecommunication Systems (CITS), Gijon, 2015, pp. 1-5. doi: 10.1109/CITS.2015.7297753

[69] Google Transparency Report "Browse real-time traffic to Google products and services" All Google products for October 14th, 2014. Internet: `http://www.google.com/transparencyreport` [March. 20, 2015].

[70] http://www.cisco.com Cisco CRS-1 16-Slot Single-Shelf System Internet: www.cisco.com/c/en/us/products/collateral/routers/crs-1- 16-slot-single-shelf-system/product data sheet09186a008022d5f3.html Jan 24, 2014 [July. 15, 2016].

[71] F. A. Ghonaim, T. E. Darcie and S. Ganti, *Impact of SDN on Optical Router Bypass* in Journal of Optical Communication and Networking, vol. 10, no. 3, pp. TBA 2018

[72] RFC 5440 - Path Computation Element (PCE) Communication Protocol (PCEP) Available: `https://tools.ietf.org/html/rfc5440`

[73] L. Gifre, A. Castro, M. Ruiz, N. Navarro and L. Velasco, "An in-operation planning tool architecture for flexgrid network re-optimization," 2014 16th International Conference on Transparent Optical Networks (ICTON), Graz, 2014, pp. 1-4.

[74] D. B. Jacobs, " Path Computation Element primer: Are PCE and SDN connected?" Available: urlsearchsdn.techtarget.com/tip/Path-Computation-Element-primer-Are-PCE-and-SDN-connected [March 23. 2017]

[75] Huawei Enterprise, Transport SDN Solution. Available: enterprise.huawei.com/ilink/enenterprise/download/HW_341919 [March 13. 2017].

[76] Scalable and efficient orchestration of Ethernet services using software-defined and flexible optical networks.

[77] R. Gewirtz Little "Making Networks Virtual: The Latest on SDN Technologies" Available: book.itep.ru/depository/open_flow/Making_Networks_Virtual_Latest_on_SDN_Technologies_hb_final.pdf 2014. [Oct. 23, 2015].

[78] Pacnet.com "PACNET FIRST TO EXTEND SDN and NFV CAPABILITIES INTO THE OPTICAL LAYER" press release. Available: www.pacnet.com/pressrelease/pacnet-first-to-extend-sdn-amp-nfv-capabilities-into-the-optical-layer 11 Mar 2015. [Oct 23. 2015].

[79] A. Thyagaturu, A. Mercian, M. McGarry, M. Reisslein and W. Kellerer,. "Software defined optical networks (SDONs): A comprehensive survey." IEEE Communications Surveys and Tutorials 18.4 (2016): 2738-2786.

[80] F. Pederzolli, M. Gerola, A. Zanardi, X. Forns, J. F. Ferran and D. Siracusa, "YAMATO: The First SDN Control Plane for Independent, Joint, and Fractional-Joint Switched SDM Optical Networks," in Journal of Lightwave Technology, vol. 35, no. 8, pp. 1335-1341, April15, 15 2017.

[81] G. Talli et al., "Multi-service SDN controlled reconfigurable long-reach optical access network," 2017 European Conference on Networks and Communications (EuCNC), Oulu, 2017, pp. 1-5.

[82] A. Sadasivarao, D. Naik, C. Liou, S. Syed and A. Sharma, "Demystifying SDN for Optical Transport Networks: Real-World Deployments and Insights," 2016 IEEE Global Communications Conference (GLOBECOM), Washington, DC, 2016, pp. 1-7.

[83] Gringeri, Steven, Nabil Bitar, and Tiejun J. Xia. "Extending software defined network principles to include optical transport." IEEE Communications Magazine 51.3 (2013): 32-40.

[84] N. Amaya et al., "Software defined networking (SDN) over space division multiplexing (SDM) optical networks: features, benefits and experimental demonstration." Optics express 22.3 (2014): 3638-3647.

[85] S. Pachnicke, H. Kagba and P. M. Krummrich, "Load Adaptive Optical-Bypassing for Reducing Core Network Energy Consumption," Photonic Networks, 12 . ITG Symposium, Leipzig, Germany, 2011, pp. 1-5. Available: http://www.ict-strauss.eu/en/ [March 13. 2017].

[86] Ciena.com "8700 Packetwave Platform". Available: : www.ciena.com/products/packet-networking

[87] Cisco.com "SDN for Service Providers". available: : www.cisco.com/web/HR/ciscoconnect/2013/pdfs/software_defined_networks_sdn_for_service_providers.pdf. [Feb. 16, 2016].

[88] E. Mannie. "Generalized multi-protocol label switching (GMPLS) architecture." Interface 501 (2004): 19.

[89] T. P. d. Souza, J. R. A. Cavalcante, A. Patel, M. E. Monteiro and J. Celestino, "SONA: Software Defined Optical Networks Slicing Architecture," 2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA), Taipei, 2017, pp. 654-661.

[90] Cisco.com "Cisco Visual Networking Index: Forecast and Methodology, 2015-2020". available: : http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html. Jun 01, 2016 [Jun. 23, 2016].

[91] Cisco.com "VNI Complete Forecast Highlights Tool: 2020 Forecast Highlights". available: : http://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html. [July. 23, 2016].

[92] Xtera.com "Xtera Announces New OTN Switch for Efficient and Resilient 100G and Beyond 100G Optical Networking" Internet: http://www.xtera.com/wp-content/uploads/2015/10/2013_02_27-OTN-Switch-Announcement.pdf. [July. 23, 2016]

[93] Cisco.com "Cisco CRS Carrier Routing System16-Slot Line Card Chassis System Description". Avialble: http://www.cisco.com/c/en/us/td/docs/routers/crs/crs1/16_slot_lc/system_description/reference/guide/qsysdsc.pdf March 2015. [July. 23, 2016]