Energy-Efficiency in Optical Networks

Saldaña Cercos, Silvia; Tafur Monroy, Idelfonso ; Fagertun, Anna Manolova

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Energy-Efficiency in Optical Networks

Silvia Saldaña Cercós

Supervisors:
Professor Idelfonso Tafur Monroy
Dr. Anna Manolova Fagertun

Delivery Date: 30th June 2015

DTU Fotonik
Department of Photonics Engineering
Technical University of Denmark
Building 343
2800 Kgs. Lyngby
DENMARK

DTU Fotonik
Department of Photonics Engineering

Abstract

This thesis expands the state-of-the-art on the complex problem of implementing energy efficient optical networks. The main contribution of this Ph.D. thesis is providing a holistic approach in a multi-layered manner where different tools are used to tackle the urgent need of both estimating and optimizing power consumption in different network segments.

An energy consumption analysis for a novel digital signal processing for signal slicing to reduce bandwidth requirements for passive optical networks is presented in this thesis. This scheme aims at re-using low bandwidth equipment to cope with current traffic demands and this dissertation tackles the trade-off between energy efficiency and quality of service in terms of latency.

Another important contribution of this thesis is the novel mixed integer linear programing (MILP) formulation for internet protocol (IP) over wavelength division multiplexing (WDM) core network design. This work provides a detailed model based on a modular architecture where the network is upgraded as the network traffic increases considering physical constraints instead of approximations on power consumption per port or assuming a single type of interface module. A comprehensive analysis on the trade-off between power consumption and capital expenditures (CAPEX) is also presented. The results confirmed that gross-grained designs (i.e., designs that account for high bandwidth technologies) are attractive from a cost perspective, however, power consumption needs to be further improved.

A second analysis for core networks is performed on a programmable networking platform, named software defined networking, to consolidate core network designs towards cost-efficient and flexible solutions. Results on jitter, latency, and power consumption for a novel south-bound protocol named KeyFlow are reported showing more than 50% reduction in the round trip time, eliminating jitter from the system, and obtaining up to

57% power reduction compared to a reference OpenFlow switch.

Furthermore, this thesis introduces an evaluation of a novel approach for short range high capacity links for datacenters. A polarization multiplexing strategy named quad-polarization, where four independent data streams are transmitted simultaneously, is evaluated from computational complexity, power consumption, and receiver sensitivity perspectives. To provide a comprehensive analysis, comparison with parallel optics and WDM systems is reported. These results show the trade-off between increased capacity and both power consumption and system performance.

In conclusion, an energy-efficient set of tools has been provided covering different aspects of the telecommunication network resulting in a cohesive research in three networks segments: access, core networks, and datacenters. Additionally, it opens prospects for next generation energy efficient networks providing tools to analyze, estimate, and optimize that new metric in optical network design: energy consumption.

Resumé

Denne afhandling bidrager med den nyeste forskning inden for energieffektive optiske netværk. Hovedbidraget er en holistisk mange-lags metode igennem hvilken forskellige værktøjer bruges til at håndtere det akutte behov for både at estimere og optimere energiforbruget i forskellige segmenter af netværket.

En analyse af energiforbruget af en ny digital signalbehandler til signal slicing, som bruges til at mindske behovet for båndbredde, er beskrevet i denne afhandling. Systemet sigter efter at genbruge lav-båndbredde udstyr til at magte nutidens efterspørgsel. Denne afhandling behandler således kompromiset mellem energieffektivitet og kvaliteten af servicen set fra et ventetids synpunkt. Et andet vigtigt bidrag fra afhandlingen er en nye mixed integer linear programing (MILP) formulering til internet protokol (IP) over wavelength division multiplexing (WDM) core network design. En detaljeret model for netværksdimensionering er givet. Denne baserer sig på en modulær arkitektur, hvor netværket opgraderes efterhånden som behovet stiger. De fysiske begrænsninger er taget i betragtning frem for at estimere energiforbrug per port eller antage én bestemt type af grænseflademodul. En grundig analyse af fordelene og ulemperne mellem energiforbruget of anlægsudgifterne er også givet. Fra resultaterne bekræftes at coarse-grained design (design som tager hensyn til høj-båndbredde teknologier) er attraktive fra et omkostning perspektiv, men energiforbruget skal forbedres.

En anden analyse af core-netværker er udført på en programmerbar netværksplatform, kaldet software defined networking, for at befæste core-netværk med omkostningseffektive og fleksible løsninger. Jitter, latency, og energiforbrugresultater fra en ny south-bound protokol kaldet KeyFlow mere end halverer latency og fjerner jitter fra systemet samt opnår en energibesparelse på 57% i forhold til en reference OpenFlow switch.

Endvidere introducerer denne afhandling en evaluering af en ny metode til at lave kort-rækkevidde forbindelser i datacentre med høj kapacitet. En polariseret multioverførsel strategi, kaldet quad-polerization, hvor fire uafhængige data-strømme overføres samtidig, er blevet evalueret ud fra beregnings kompleksitet, energiforbrug, og modtagerfølsomhed. En sammenligning mellem parallel optik og WDM systemer er udført for at levere en grundig analyse. Her bliver den øgede kapacitet afvejet mod både energiforbrug og systemets ydelse.

Afslutningsvist er et sæt af værktøjer til energieffektive løsninger leveret. Værktøjerne dækker forskellige grene af telekommunikationsnetværkene, der resulterer i sammenhængene forskning inden for 3 netværkssegmenter: access, core-netværk, og datacentre. Desuden baner det vejen for næste generation af energieffektive netværk ved at levere de nødvendige værktøjer til at analysere, estimere og optimere denne nye udfordring i design af optiske netværk: Energieffektivitet.

Acknowledgements

First I would like to thank my supervisors, Professor Idelfonso Tafur Monroy, and Dr. Anna Manolova Fagertun. Idelfonso for providing me with the opportunity to start my research career at DTU. Anna, for all the time that she has invested in me. She showed me the importance of being methodical and organized at work.

I would also like to thank Professor Moisés R. N. Ribeiro for being an invaluable guide throughout this project. He served as inspiration for hard work with both passion and discipline. Many thanks Moisés for your precise comments, good advice, priceless supervision and help. I also would like to thank you for being my supervisor, neighbor, and family during my external stay.

I would like to thank all the co-authors for making this work possible. I thank all my colleagues at the metro-access group, and also my colleagues in DTU Fotonik.

Additionally, I would like to thank my salsa dancing friends and my basketball teammates for providing me with the energy to clean my mind and start every day fresh with new ideas.

I would like to thank my family, for giving me support during this time. My family who has always been there, for more than 6 years (even though the initial deal was 6 months).

Last, I would like to thank the person who made all this possible, a person that has kept me alive during the entire process: Stig Hoegberg. Thanks for believing in me, in what I do, and for all the support from the first day we met.

Summary of Original Work

This thesis is based on the following original publications:

PAPER 1

S. Saldaña Cercós, C. Wagner, J. J. Vegas Olmos, A. Manolova Fagertun, and I. Tafur Monroy, "Digital signal processing for a sliceable transceiver for optical access networks," *Proc. 20th symposium on computers and communications, ISCC'15*, Larnaca, Cyprus, 2015.

PAPER 2

S. Saldaña Cercós, L. C. Resendo, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, "Power-aware multi-layer translucent network design: an integrated OPEX/CAPEX analysis," in *Proc. Optical Fiber Communication Conference and Exposition, OFC/NFOEC*, San Francisco, California, USA, 2014.

PAPER 3

S. Saldaña Cercós, L. C. Resendo, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, "A power-aware rationale for using gross-grained transponders in IP-over-WDM networks," *Journal of optical communications and networking*, under review, 2015.

PAPER 4

S. Saldaña Cercós, G. E. Rordrigues de Paiva, M. Colazza Argentato, A. Manolova Fagertun, J. Rodrigues Fernandes de Oliveira, and I. Tafur Monroy, "Empirical multichannel power consumption model for erbium-doped fiber amplifiers," *In submission.*

PAPER 5

S Saldaña Cercós, R. M. Ramos, A. C. Ewald Eller, M. Martinello, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, "Design of a stateless low-latency router architecture for green software-defined networking," *Proc. Photonics West SPIE'15*, San Francisco, California, USA, 2015.

PAPER 6

S. Saldanã Cercós, R. E. Oliveira, R. Vitoi, M. Martinello, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, "Tackling OpenFlow power hog in core networks with KeyFlow," in *Electronic letters*, **50**(24), 1847-1849, 2014.

PAPER 7

S. Saldanã Cercós, M. Piels, J. Estarán, M. usuga, E. Porto da Silva, A. Manolova Fagertun, and I. Tafur Monroy, "100 Gbps IM/DD links using quad-polarization: Performance, complexity, and power dissipation," *Optics Express*, under review 2015.

Other scientific reports associated with the project:

PAPER A

I. Tafur Monroy, A. Caballero, **S. Saldanã Cercós**, and R. Borkowski, "Cognition-enabling techniques in heterogeneous and flexgrid optical communication networks," in *Proc. Photonics West SPIE'12*, 2012.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and problem statement

The number of Internet users in the world is increasing with an annually growth rate of 8% [1]. In June 2014 more than 87% and 70% of the North American and European population, respectively were Internet users. Accounting for the world total population, more than 3 billion users are registered as Internet users which represents more than 40% of the world population [2]. Additionally, the emerging variety of internet access technologies have lead towards a new lifestyle, where heterogeneous and dynamic services such as online social networks (OSNs), online gaming, video streaming on-demand, e-health, and/or e-commerce are present in end-users' daily lives, pushing further the traffic demands with a forecasted IP traffic growth of 21% for the next five years [3]. This traffic growth leads to a corresponding increase in power consumption. The Information, communication, and technology (ICT) sector's CO_2 emissions are expected to be $2-3\%$ of the total emissions by 2020 [4].

The ICT sector's power consumption is a concern not only from primary resources scarcity and environmental impact point of view, but also due to electricity generation fundamental limitations. According to [5], if there is no technology development, energy consumption will increase proportionally to the traffic growth, and as shown in Fig. 1.1, the electricity required only for network equipment (NE) and datacenter (DC) will reach the total electricity generation by 2022.

There is an unmet need to improve the current Internet infrastructure to leap the barrier of providing the resources to cope with high traffic demands in an energy-efficient manner. Future trends indicate that taxation strategies to encourage efficient utilization of electricity are under consideration [6]. Thus, along with CO_2 emissions and resources scarcity, energy efficiency in optical networks has a socioeconomic impact.

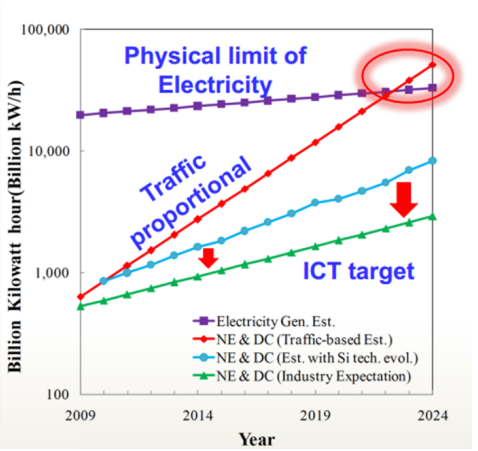


Figure 1.1: World-wide electricity generation and consumption. DC, datacenter; NE, network equipment [5]

Problem statement

Optical network designers face the challenge of providing enough capacity to cater for the aforementioned traffic growth but network upgrade solutions cannot be detached from energy efficiency considerations. However, the strong competition in the telecommunication market motivates network providers to satisfy traffic demands in an energy-effective manner at no Quality of service (QoS) nor cost expenses. Thus, network designers seek for power consumption models of the Internet infrastructure and analyses of factors that influence the overall power consumption to tackle trade offs between green networking, cost, and QoS.

The focus of this Ph.D. project is to provide the community with input within this field by evaluating power consumption in three different network segments tackling relevant trade offs such as power consumption with increased capacity, cost, and/or system performance either in terms of latency, or receiver sensitivity.

1.2 Methodology

The work presented in this dissertation is based on an interdisciplinary methodology as illustrated in Fig. 1.2. The synergy between mathematical formulations (i.e., MILP), mathematical evaluation, experimental work, and hardware implementation results in a multi layered contribution covering the complex problem of energy consumption in optical networks from system level to architecture level both tackling hardware and software implementations. This combined research both at the physical layer and at the control plane can overcome the barrier of presenting a complete overview of energy efficiency in optical networks. **PAPER 1** and **PAPER 7** are based on a post-processing analysis based on a mathematical evaluation of a given software. Experimental work at the physical layer has also been assessed in the later contribution as well as in **PAPER 4**. **PAPER 2** and **PAPER 3** are based on MILP formulations for network design. Finally, **PAPER 5** and **PAPER 6** are based on NetFPGA-based implementations, where power consumption analysis have been performed using a design software tool, Xilinx XPower Analyzer (XPA), after synthesizing the design.

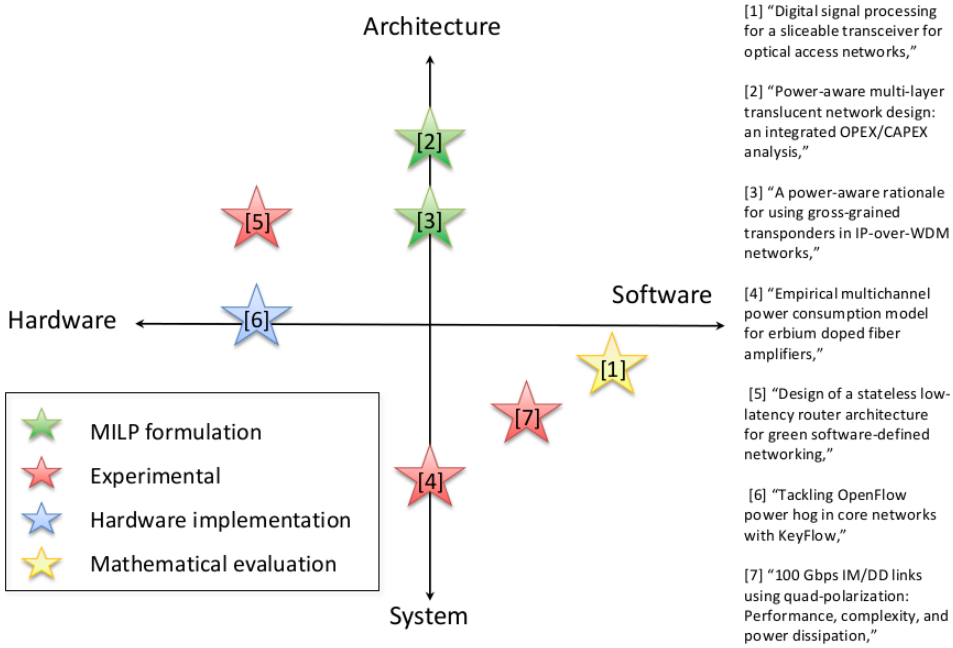


Figure 1.2: Ph.D. project contribution

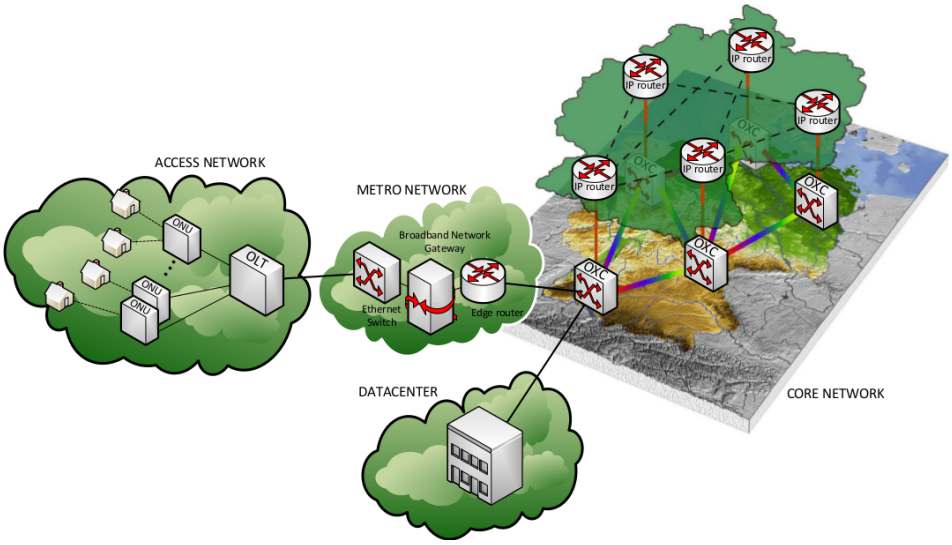


Figure 1.3: A high-level network structure with four different segments: Core network, metro network, access network, and datacenters. IP, Internet protocol; OLT, optical line terminal; ONU, optical network unit; OXC, optical cross connect

1.3 Network segments

A high-level overview of a standard telecommunication network infrastructure is provided in Fig. 1.3. The network is split in so-called network segments. This network fragmentation eases the design of a platform to model the telecommunication infrastructure power consumption, where the challenges and trade offs of each segment can be tackled individually.

The network architecture is composed of three segments which are access, metro, and core networks. Access networks connect the end users to the local exchange from the metro network. Metro networks connect local exchanges to each other, aggregate fluctuating traffic from the end user, and act as an interface between the access and the core networks. Finally, core networks which are composed of a small number of large (i.e., high capacity) routers, transport high bandwidth data streams through long distances interconnecting geographically separated locations.

Additionally, datacenters are considered important power consumption contributors within the network infrastructure. They took an important role from the late '90s with the dot-com bubble and have become focus of interest from a power consumption point of view.

This thesis targets three of these four main network segments. The contributions to push the state of the art forward in access networks are described in Chapter 2. Rresearch carried on core networks is summarized in Chapter 3 which includes both current well-established technologies such as mixed line rate (MLR) wavelength division multiplexing (WDM) architectures and new trends such as software defined networking (SDN). Finally, contributions associated to datacenters is detailed are Chapter 4.

The following subsections present an overview of each of the network segments, its associated challenges, requirements, and trade-offs.

1.3.1 Access networks: passive optical networks

Next generation access networks need to cope with the challenge of providing high data transmission rates (which have moved from hundreds of Mbps up to multi-Gbps) in an heterogeneous environment where both wired and wireless technologies take place. In terms of power consumption, the access network is currently the most power demanding segment of the Internet and is expected to remain such in the medium term future [7], therefore an active research area. The power consumption in access networks is proportional to the traffic generated by end users. Even though it could be considered that all traffic generated in access networks is further aggregated in core networks, some studies work on content distribution, where the main goal consists of placing the data closer to end user, thereby sparing transportation in back-haul networks. Initially, content distribution networks (CDNs) were designed to improve the end user QoS in terms of latency though later on CDNs were also exploited to reduce power consumption [8].

Fiber based wired access networks (e.g., fiber to the home (FFTH)) were introduced as an alternative to copper-based wired access networks (i.e., digital subscriber line (DSL)) since they offer increased data transmission rates due to the large bandwidth provided by the optical fiber [9]. Passive optical networks (PONs) are an attractive solution to implement FFTH networks because they offer high bandwidth capacity as well as low power consumption due to the use of passive components on the signal's path. Additionally, access networks are also cost sensitive, and thus cost effective solutions are currently focus of research [10].

PAPER 1 introduces a new digital signal processing (DSP) for signal slicing to reduce bandwidth requirements for PONs while updating network resources by re-using low bandwidth equipment to cope with current traffic

growth. Signal slicing in PON access networks aim at decreasing power consumption of the end-to-end access network, as well as at facilitating the migration from low to high bandwidth technologies.

How to evaluate any DSP power consumption is described in Section 1.4, where calculations are based on the number of operations (i.e., real multiplications and real additions) required by the DSP.

1.3.2 Metro networks

Metro networks act as an interface between access networks and core networks. They are composed of three elements: 1) Ethernet switches, 2) broadband network gateways (BNGs), and 3) edge routers. The Ethernet switches aggregate the traffic from a large number of access network nodes. They are connected to BNGs which are responsible of performing access rate control. They enforce QoS policies and provide layer 3 (i.e., IP layer) connectivity to route traffic through the edge routers. The edge routers connect to the core network.

According to Alcatel-Lucent metro network has been under pressure since the hot spot in terms of traffic growth is shifting from access to metro networks [11]. A new architecture where metro and core networks are merged has been proposed in [11] to address the challenges that arose with cloud and mobile broadband services.

1.3.3 Core networks

In long term future, power consumption will be dominant in the core network as the bitrates increase [7]. Consequently, part of this thesis turns its attention to core networks. Two different scenarios are considered. First, a two-layer network model is studied. Internet protocol (IP) over WDM, is considered as a network architecture since it aims at reducing network components (i.e, shelves, processors, interfaces cards) by facilitating integration, and has been proven to be an energy-efficient solution for optical backbone networks [12]. Second, a programmable networking platform, named SDN, is evaluated as a strategy to consolidate core network designs towards cost-efficient and flexible solutions.

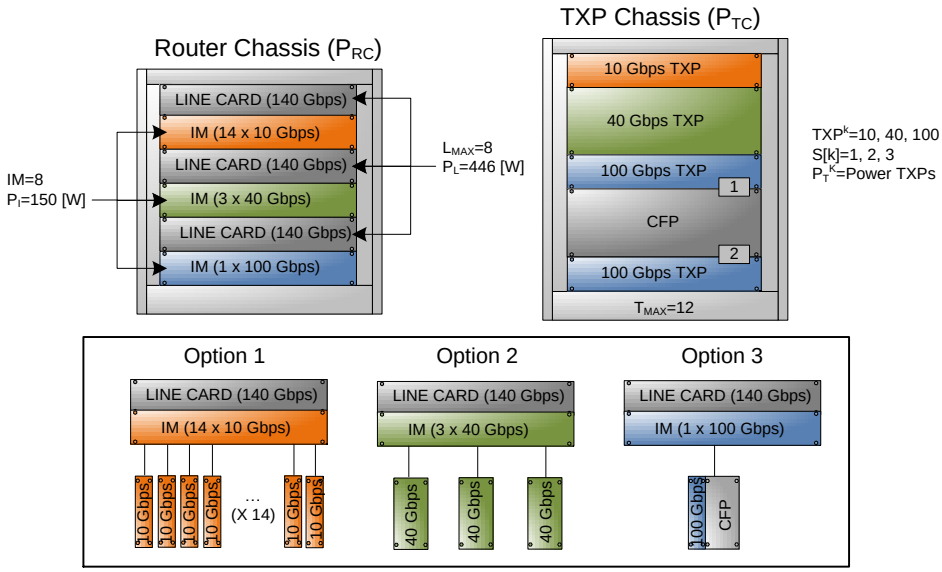


Figure 1.4: IP-over-WDM modules. CFP, C form-factor pluggable; IM, interface module; IP, internet protocol; TXP, transponder; WDM, wavelength division multiplexing

IP-over-WDM core networks

The two-layer model used for evaluating IP-over-WDM power-aware designs is based on a modular configuration. The main advantage of a modular configuration is that scalability is achieved by interconnecting extra modules per node instead of changing the entire infrastructure. The top left of Fig. 1.4 depicts the modular configuration based on a single chassis configuration from Cisco CRS-3 router, which has been chosen due to the power consumption information availability and because Cisco technology is worldwide deployed. Each router is composed of a number of router chassis which provide 16 slots for a diversity of line cards which provide 140 Gbps throughput, and an interface module (IM) for each line card. The IMs can provide either 1, 3, or 14 interfaces at 100, 40, or 10 Gbps, respectively. The IM are used as interface between the IP layer and the WDM layer.

The top right of Fig. 1.4 illustrates the WDM layer modular configuration. At the WDM layer, a transponder chassis with 12 slots capacity houses a number of transponders with different line rates (i.e., 10, 40, or 100 Gbps). Note that for this particular configuration, an additional C-

form factor pluggable (CFP) is required for the 100 Gbps technology.

The IP layer and the WDM layer are interconnected as shown in Fig. 1.4 bottom with correspondence 1:1 between transponders an interface modules.

In a multi-layer networking architecture there are three different strategies to deal with traffic. Fig. 1.5 illustrates the possible paths from source to destination. In an opaque network (path a), there is optical-electrical-optical (OEO) conversion at each intermediate node. In a transparent network, a direct link that by-passes each intermediate node is used (path b). Translucent networks perform OEO-conversion at intermediate nodes only when the network can benefit from it or it is strictly necessary (e.g., to aggregate traffic at the intermediate node (i.e., grooming) or to perform signal regeneration). Translucent networks have been demonstrated to be the most cost-effective and energy-efficient [13], thus for the optimization models presented in this thesis, only translucent networks are considered.

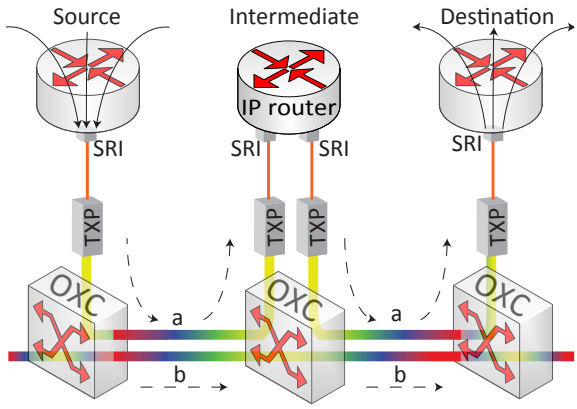


Figure 1.5: Multi-layer networking strategies. Path a) opaque, b) transparent. OXC, optical cross connect; SRI, short range interface; TXP, transponder

PAPER 2 presents a combined capital expenditures (CAPEX)/operational expenditures (OPEX) analysis for IP-over-WDM translucent networks with modular architecture. It presents a model based on mixed integer linear programming (MILP) to reduce OPEX by minimizing the number of network reconfigurations and decreasing the overall power consumption.

PAPER 3 presents an extension of the model to analyze the power consumption of the network of the future after adoption of 40 Gbps and 100 Gbps technologies.

Regarding the link power consumption for the erbium-doped fiber amplifier (EDFA) contribution, models that only consider a static power consumption contribution are found in literature [14]. In the two models presented in **PAPER 2** and **PAPER 3** the EDFA power consumption is consider fixed, or channel-independent, thus link power consumption for the optimization model is assumed to be background power consumption, therefore not accounted for. This work is extended in **PAPER 4** by presenting a power consumption model of EDFAs based on empirical data which accounts for channel dependency.

SDN for core forwarding engines

Additional to the traditional multi-layer core network architecture, in this thesis a more flexible implementation based on SDN is studied. The main goal for network providers is to provide cost-efficient solutions by implementing low number of core network nodes which support high bitrates. One approach to reduce cost is to design core network forwarding engines that can support intra-domain routing capabilities while mitigating the node complexity. The reference south-bound protocol for SDN, OpenFlow, does not provide the properties to offer low-latency, low-complexity, and energy-efficient solutions while assuring network scalability [15].

PAPER 5 and **PAPER 6** introduce a new south-bound protocol named KeyFlow which eliminates the need for tables to keep track of the states associated to active flows. KeyFlow replaces flow tables by a simple operation reducing core node complexity. Latency, jitter, and power consumption analysis are presented in Chapter 3.

1.3.4 Data centers

The last section, or segment, that this thesis puts focus on is the data center. With the increased number of applications that are based on cloud services, optical data links are suffering from bandwidth limitations. Short range links aiming data centers' infrastructures are hot topic for research both at the industry and the academia levels.

Data center's challenges include among others increasing optical bandwidth while dealing with space constraints. Small footprint networks with high density equipment yield to an increase in data center's electricity use. According to [16] in 2010 data center's electricity already accounted for

1.3% of the world electricity use, and 2% of all electricity use for the US. Consequently, any new strategy to increase optical bandwidth needs to be evaluated from a power consumption point of view before its implementation.

PAPER 7 provides a comparison between three different strategies to increase bandwidth: wavelength division multiplexing (WDM), parallel optics, and a novel polarization multiplexing solution proposed in [17] referred throughout this thesis as quad-polarization. The comparison encloses computational complexity, power dissipation, and receiver sensitivity.

1.4 Digital signal processing power consumption

This section provides the tools to evaluate the power consumption of any DSP subsystem for any segment in the network. The methodology used throughout this thesis to evaluate the power consumption for a specific DSP subsystem is based on the number of real operations (i.e., real multiplications and real additions) and the power consumption associated to each of them based on the applications-specific integrated circuit (ASIC) used to implement it.

This method is used for calculating the power consumption of the additional DSP required for the sliceable transceiver presented in **PAPER 1** and of the needed DSP for the polarization tracking algorithm for short range links presented in **PAPER 7**. Both papers are based on a proof-of-concept where no power consumption optimization had been performed. Also, the power consumption associated to the DSP depends on the implementation (i.e., ASIC, NetFPGA, FPGA...). Consequently, the method outlined herein is design to provide an overview and upper boundary of each DSP subsystem power consumption rather than give exact and optimized estimates.

The results presented in this thesis are based on a 90 nm, CMOS process technology. A commercially available ASIC design from Nortel is used for the numerical calculations since it is well documented with detailed information on its power consumption. It consumes 1.5 pJ and 0.5 pJ per real multiplication and real addition, respectively [18].

Let E_m and E_a be the energy in Joules that each real multiplication and each real addition consumes, respectively. The total energy consumed by any DSP subsystem can be calculated as indicated in equation 1.1

$$E_i = Q_i \cdot E_m + A_i \cdot E_a \tag{1.1}$$

where E_i is the energy consumption of the DSP subsystem i (e.g., fast fourier transform (FFT), rectifying filter, polarization tracking algorithm, etc...), Q_i and A_i stand for the total number of multiplications and additions required for the DSP subsystem i, respectively.

Chapters 2 and 4 provide a detailed description of each DSP subsystem (E_i) for sliceable transceiver and quad-polarization strategy, respectively, breaking them down into number of real operations to evaluate the associated DSP power consumption.

1.5 Main contribution and outline of the thesis

This thesis is based on a set of articles already published or submitted for publication in peer-reviewed journals and conference proceedings. These articles present the results obtained during the course of my doctoral studies into the design, analysis, and evaluation of different strategies to achieve energy efficient optical networks, accounting for solutions in access and core networks, as well as in datacenters. The work presented here combines analysis, simulation and experimental work.

The papers are grouped in three categories based on the network segment the proposed solution is to be implemented in. **PAPER 1** deals with signal slicing for access networks. Core network design based on IP-over-WDM architectures is covered in **PAPER 2** to **PAPER 4**. An extension for core networks is presented in **PAPER 5** and **PAPER 6**, where the benefits of a flexible programmable platform (i.e., software defined networking) are exploited for core network fabrics. Finally, **PAPER 7** presents an analysis and comparison of three different solutions for short range intensity modulated/direct detection (IM/DD) links for datacenter environments.

PAPER 1 presents signal slicing as a method to upgrade the network infrastructure to cope with current traffic demands. The state of the art in access network architectures is pushed forward by proposing a solution using signal slicing to re-use low bandwidth equipment to satisfy high bandwidth traffic demands. Signal slicing is based on DSP, which comes at the cost of increased processing delay. In this paper a survey of the required DSP subsystems is reported, including a detail description of each block,

the power consumption associated to it, and latency penalties. The power consumption analysis is based on the number of operations (i.e., real additions and real multiplications) required by the DSP. The methodology used to calculate the associated power consumption is described in Section 1.4. Results are discussed for both WDM and time division multiplexing (TDM) solutions, at both 1 Gbps and 10 Gbps. Experimental validation for the 1 Gbps scenario is presented, whereas simulations are used for higher bitrates (i.e., 10 Gbps). For 1 Gbps WDM system 278 pJ per information bit for 4 slices is reported at 105 ns latency penalties, whereas 3898.4 pJ per information bit at 183.5 μs latency penalty is reported for 10 Gbps. Power savings in the order of hundreds of Watts can be obtained when using signal slicing as an alternative to currently implemented 10 Gbps access networks.

PAPER 2 reports on a novel multi-phase model which provides a design for IP-over-WDM core networks and an approach for traffic engineering accounting for daily traffic dynamics. The main contribution of this work is the definition of an optimization model which accounts both for CAPEX at the design phase, and OPEX in terms of power consumption and reconfiguration at the traffic engineering phases. The proposed model is based on MILP formulation which considers a modular architecture (i.e., the network modules such as router line cards or optical transponders are upgraded as the network traffic increases). The model has been created to compare a cost-effective design versus a power-aware design. Network design based on power consumption yields to only 6% increase in CAPEX. Additionally, a minimal 4.2% power consumption increase is achieved when constraining network reconfiguration, thus, increasing QoS since reconfiguration requires service disruption when traffic needs to be re-route while applying traffic engineering. This paper captures the design trade-offs and makes a compelling argument for designing towards energy efficient optical networks.

PAPER 3 presents an extension of the model proposed for network design in **PAPER 2**. The model is used to define the impact in terms of power consumption of employing 100 Gbps optical transmission technology to cope with upcoming traffic demands. State-of-the-art MILP models for network design offer solutions to over-provision the network in order to cope with peak traffic demands without considering its OPEX impact in terms of power consumption. In this paper solid results on power consumption of the network of the future after the adoption of 40 and 100 Gbps technology are presented. Results show that 37.7% optical network capacity increase

is achieved when using exclusively 100 Gbps at 13.3% to 32.4% power consumption expenses. In case of using a design combining 40 and 100 Gbps technologies, 15.3% optical network capacity increase is achieved at only 2.6% to 8.8% power consumption penalty.

PAPER 4 presents the first known experimental power consumption analysis and model of single and multi-stage booster EDFAs accounting for channel dependency. We have expanded the state of the art by considering a dynamic power consumption contribution for EDFAs, since models found in literature consider EDFA power consumption constant and independent on the number of simultaneously amplified channels. Assuming no power control on the EDFA pump laser, limits the achievable power savings for backbone networks. Results show that 48% of the total power consumption is associated to EDFA channels loading.

PAPER 5 and **PAPER 6** consider a flexible programming platform for core networks. Software design networking for control of transport networks presents scalability challenges since it has to deal with the large number of flows that core network fabrics have to handle. A south-bound protocol name Keyflow is analyzed to offer a cost- and energy-efficient forwarding fabric for core networks. The work presented here includes latency, jitter, and power consumption analysis. Results show that the packet round trip time (RTT), thereby latency, can be reduced more than 50% using KeyFlow instead of the reference protocol OpenFlow. Jitter reducion is demonstrated experimentally, and power consumption savings above 57% are achieved.

In **PAPER 7** a comparison of three different strategies for short range high capacity links for datacenters is presented. The scenarios considered are quad-polarization, WDM, and parallel optics. In order to evaluate each scenario's benefits, the analysis includes a comparison of computational complexity, power dissipation, and QoS in terms of receiver sensitivity. A negligible 2.47% additional power consumption is required by the quad-polarization DSP compared to WDM and parallel optics due to its simplicity. However, a 47.6% power consumption increase is experienced due to the required analog to digital converters (ADCs). In terms of receiver sensitivity, -10.5 dBm for 7% forward error correction (FEC) overhead are achieved for parallel optics with 1 dBm penalty for the WDM system due to the arrayed waveguide grating (AWG) insertion loss. For the quad-polarization system, 4.4 dBm receiver sensitivity is obtained with 0.5 dBm penalty after 2 km of standard single mode fiber (SSMF) transmission.

Chapter 2

Optical access networks

2.1 Sliceable transceiver for access networks

Sliceable transceivers are an alternative for a capacity upgrade strategy based on the re-use of low bandwidth equipment. The main idea of sliceable transceivers consists in breaking a high bandwidth signal into smaller slices to be transmitted either in parallel benefiting from WDM technology, or in serial using TDM technology. The algorithm used for signal slicing has been used and implemented in real time oscilloscopes to provide with full signal path symmetry which offers lower noise and jitter, leading to higher bandwidth oscilloscopes (i.e., 70 GHz) [19]. Additionally, signal slicing can be used to enhance security. This extra level of security can be achieved by transmitting each of the signal slices through different transmission media (e.g., optical, copper, or wireless links).

Re-using lower bandwidth equipment brings several benefits such as re-using already deployed equipment, offering network providers the possibility to reduce CAPEX by extending the lifetime of current infrastructures and still coping with current traffic demands. Additionally, low bandwidth equipment overcomes the disadvantage of over-provisioning the network with high bandwidth and costly equipment to satisfy peak traffic demands while under-utilizing the resources in the low traffic hours (e.g., during the night).

New strategies to cope with bandwidth limitations in optical networks need to be also evaluated from a power consumption point of view. Signal slicing has the potential to provide power savings since even though high

bandwidth technology is more power efficient (i.e., W/Gbps) [20], for traffic loads below the equipment capabilities it can represent a power expense. Additionally, lower bandwidth equipment might eases the use of on/off strategies since there is no need for high bandwidth and power demanding equipment to satisfy low traffic demands [21].

In this section the DSP required for signal slicing is explained. Discussion and results are summarized in Section 2.3. **PAPER 1** presents a detailed description of the DSP, its computational complexity, and evaluates the trade-off between power consumption and latency. The methodology used to evaluate the DSP power consumption is the one described in Section 1.4.

2.2 Digital signal processing for signal slicing

Fig. 2.1 illustrates the two aforementioned strategies for implementing signal slicing for downlink in optical access networks: a) using TDM technology, and b) using WDM technology.

DSP is required both at the transmitter (i.e., optical line terminal (OLT)) and at the receiver (i.e., optical network unit (ONU)). At the transmitter four different DSP blocks are required in case of using WDM for the optical distribution network (ODN), and one additional block is required if TDM systems are used. DSP is performed in the frequency domain, thus the first block consists on a fast fourier transform (FFT) to

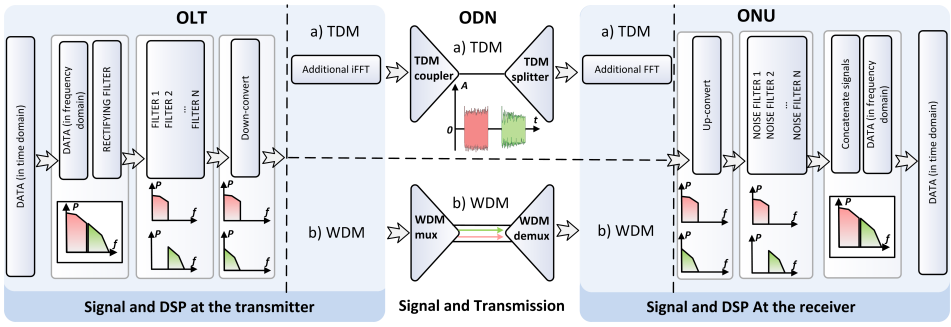


Figure 2.1: DSP for signal slicing. DSP, digital signal processing; FFT, fast Fourier transform; iFFT, inverse fast Fourier transform; OLT, optical line terminal; ODN, optical distribution network; ONU, optical network unit; TDM, time division multiplexing; WDM, wavelength division multiplexing

bring the signal from time to frequency domain. Due to the FFT mirrored sidebands appear in the signal. These are suppressed by using a rectifying filter. In order to achieve n number of slices, n pass band, filters are used. The final step at the transmitter side consists on bringing each slice to baseband. This down-conversion enables the use of equipment that mitigates the bandwidth requirements by a factor n.

At the ODN two transmission approaches have been evaluated. First, the use of TDM legacy channels. In this case, an additional inverse fast Fourier transform (iFFT) is required to bring the signal back to the time domain, and additional FFT at the receiver. Experimental demonstration of signal slicing using TDM transmission is reported in [22]. The drawbacks of TDM is that it is time sensitive, and it requires scheduling techniques and synchronization since two slices cannot co-exist. WDM enables parallel transmission using different optical carriers, enhancing bandwidth capacity.

The transmitter complementary DSP blocks are implemented at the receiver. As illustrated in Fig. 2.1, each baseband signal is brought back to the original frequency band. Then, n pass band filters are used to eliminate the amplified spontaneous emission (ASE) noise added during transmission. Finally, the full wide bandwidth signal is recovered by concatenating each of the slices and transformed back to the time domain.

2.3 Results and discussion

The design of sliceable transceivers and the associated DSP have been demonstrated experimentally for 1 Gbps non-return to zero on-off keying (NRZ-OOK) in [22] , and for 56 Gbps NRZ-OOK in [23]. Table 2.1 gathers the main results reported on sliceable transceivers.

Table 2.1: State-of-the-art for sliceable transceivers for access optical networks

Bitrate [Gbps]	Modulation format	Num. slices	Methodology
1	NRZ-OOK	2	experimental [22]
10	NRZ-OOK	4	simulation [**PAPER 1**]
56	NRZ-OOK	6	experimental [23]

Sliceable optical transceivers are a hot topic in telecom systems, though their applicability for access networks which are more critical in terms of

energy-efficiency was not determined. The work presented in **PAPER 1** extends the State-of-the-art by quantifying the added complexity and power consumption, as well as the latency characterization that is associated with the corresponding digital signal processing required for signal slicing for both 1 Gbps and 10 Gbps. Additionally, different design trade-offs are evaluated. Low bandwidth equipment is more affordable and requires less power consumption, however it presents higher computation complexity for the DSP, leading to higher latency. This is an important aspect to tackle, as the access optical networks are the most cost- and energy-sensitive of all optical segments.

Results enclose characterization for TDM systems to evaluate the plausible benefits of sliceable transceivers as an alternative solution to 10 Gbps TDM PON networks, as well as characterization for WDM systems, which boost upgrading TDM optical access networks to next generation access networks. For 1 Gbps signal WDM outperforms TDM systems both from power consumption and latency perspectives. WDM offers between 19.43% and 69.54% power consumption reduction depending on which sampling frequency is used, as well as up to 65% lower latency compared to TDM. For 10 Gbps both WDM and TDM present similar performance. This result is expected because for higher bit rates, the additional computational complexity associated to the FFT and iFFT blocks for transmission in the time domain is negligible compared to the computational complexity associated to the signal slicing itself.

A broader overview is also provided by presenting one possible implementation for the complete system. The exact power consumption for the complete system is implementation-dependent, however, results show the potential power savings that can be obtained by mitigating the bandwidth requirements of electrical and optical equipment when using sliceable transceivers. While the additional DSP contributes with hundreds of miliWatts to the overall power consumption, power savings from electrical and optical components are from few tens of Watts up to hundreds of Watts depending on the bitrate.

This thesis has contributed to the State-of-the art by presenting a thorough analysis of the DSP making a compelling argument for using signal slicing systems for optical access networks.

Chapter 3

Core network design strategies

3.1 Multi-layer translucent core network design

There are two trends for backbone network design: IP-over-(D)WDM (i.e., dense wavelength division multiplexing), and IP over optical transport network (OTN) over WDM. There is no consensus on which of the two architectures is the way to go. On one hand, IP over OTN is demonstrated to be the most cost-efficient solution according to [24] and [25]. On the other hand, IP over WDM is demonstrated to be the most energy-efficient solution [12]. In [26] it is shown that the most cost-efficient solution is not always the least energy consuming. The question of whether is better to add IP features to the OTN layer or transport feature to the router is not discussed in this thesis. In this thesis a solution to achieve integration and efficient consolidation of network layers is assumed, because even though the OTN layer outperforms the IP layer while handling TDM traffic, most of the traffic nowadays is IP traffic. Thus, IP-over-WDM is used for the design optimization models.

The two-layer model from Fig. 3.1 is used for CAPEX and OPEX analysis of backbone networks, where the optical layer is represented by the optical cross connect (OXC), and the electrical layer on top by the IP routers. Further description of this configuration is presented in Section 1.3.3.

Network providers aim at coping with current traffic demands at the lowest cost possible. MLR network architecture is a cost-effective solution to

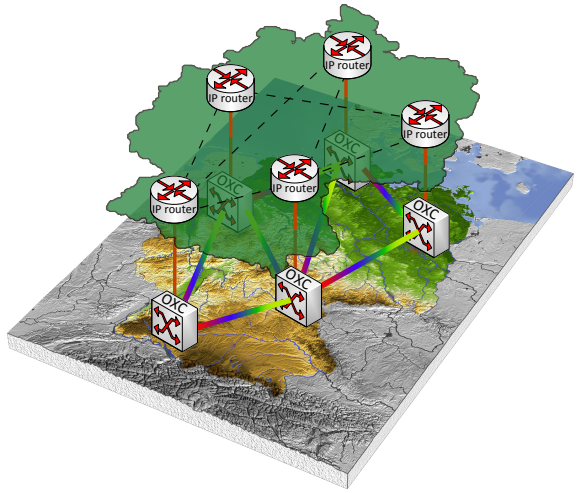


Figure 3.1: IP-over-WDM model- Optical switching at OXC and IP traffic processing at the IP router. OXC, optical cross connect; IP, internet protocol.

satisfy heterogeneous traffic demands [27]. The work reported in this thesis is based on a traditional, mature, and well-established WDM technology, considering the use of flexible transceivers as an appealing alternative to MLR configurations.

There are two phases where network providers can focus on in order to reduce cost: 1) The design phase, where which equipment is to be deployed needs to be evaluated, and 2) the traffic engineering phase, where different strategies to cope with current traffic dynamism are applied using the resources already deployed in the design phase. In this thesis both phases are assessed defining a cost-aware and a power-aware MILP formulation. Detailed information on the MILP formulations are specified in **PAPER 2** and **PAPER 3**.

Regarding the design phase, two different approaches are analyzed in **PAPER 2**. First, a cost-aware MILP formulation is used to minimize CAPEX based on normalized equipment cost (using the 10 Gbps transponder as a normalization factor), and second a power-aware MILP formulation is used to minimize OPEX in terms power consumption, where power consumption values of the modular configuration described in Section 1.3.3 are taken into account.

Once the equipment is deployed, network operators need to deal with the network traffic dynamism which is dictated by the end users' behavior.

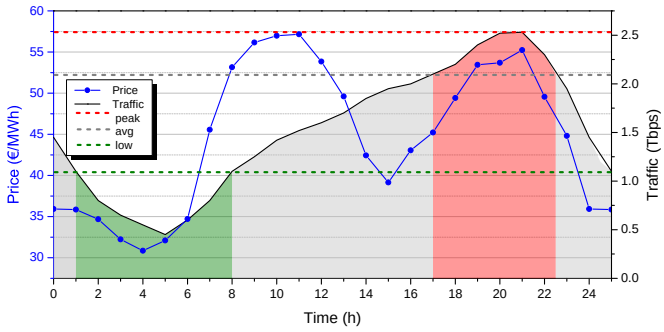


Figure 3.2: Daily traffic variations (solid back) and price (circle blue).

This behavior is changing due to the technology evolution and new trends (e.g., smart-phones, smart-cities...), yet there is one parameter that is not going to vary. This parameter is known as the biological constant. The biological constant stands for the biological behavior of humans (i.e., being awake during the day and sleeping an average of 8 h during the night). This constant defines a pattern on traffic flows where the average traffic accounts for 80% of the peak traffic, and the minimum traffic accounts for 40% of the peak traffic [28]. Fig. 3.2 illustrates this traffic variation for one of the nodes of the Deutsche Telekom (DT) network from Germany (solid black) and the electricity pricing in Germany (solid blue) for 24 h. Thus, in phase two (i.e., traffic engineering phase) a subset network architecture from the first design-phase is determined to reduce OPEX while still meeting traffic demands.

Based on the statistics presented in [29], OPEX is composed of 6 different components: certification upgrades, network care cost, space cost, cooling, services cost, and power consumption cost. In the work presented in this thesis two of these components are considered to reduce OPEX: power consumption and service cost, where service cost has been defined as the number of network reconfigurations. Consequently, two different traffic engineering strategies are evaluated. First, reducing power consumption allowing reconfiguration, and second reducing power consumption not allowing reconfiguration to avoid service disruption when resting the network to the optimal traffic engineering state, and thus reducing QoS [30].

For the integrated OPEX/CAPEX analysis presented in **PAPER 2**, two design strategies and two traffic engineering approaches are evaluated from CAPEX and OPEX cost perspectives. The two design strategies

(i.e., define which equipment is required at each node to satisfy traffic demands)include network design while minimizing OPEX in terms of power consumption, and while minimizing CAPEX. The two traffic engineering approaches include defining the routing strategy while minimizing power consumption with and without reconfiguration allowed. In order to evaluate the OPEX associated to the power consumption, the cost is estimated by multiplying the MWh consumed by every network configuration (i.e., from Fig. 3.2 low (green), average (grey), and peak (read) traffic) with the price at that specific time (solid blue line in Fig. 3.2).

3.1.1 Results and discussion

The cost-aware MILP model to define a network design aimed to minimize cost yields to only 6% CAPEX savings compared to a network design aimed to minimize power consumption. This result shows the potential of working on a design aimed at reducing power consumption due to taxation future trends. As an example, the Philippine government had already implemented a energy tax on electricity power consumption for residential customers to promote the efficient utilization of electricity in 1979 [6]. Similar strategies can be applied in the future for big entities such as network providers, thus encouraging power-aware network designs to avoid heavy tax burden.

From traffic engineering point of view, increasing QoS by avoiding service disruption when reconfiguring the network to utilize the minimum amount of resources to meet lower traffic demands, introduces only 4.2% power consumption increase. It is assumed that by reducing network reconfiguration, significant overall OPEX savings can be achieved due to service cost reduction, which accounts for more then 58% of the network OPEX.

3.2 Gross-grained design

The second analysis presented in this thesis based on the two-layer design MILP formulation, extends the previous work by evaluating the impact of gross-grained designs (i.e., imposing the use of high bandwidth technologies such as 40 Gbps and 100 Gbps) compared to the legacy fined grained-designs (i.e., using 10 Gbps technology). MLR networks have evolved from 2.5 Gbps, to 10 Gbps, 40 Gbps, and currently 100 Gbps rates. 100 Gbps

optical transmission technology is considered a plausible solution to cope with nowadays traffic demand increase, but with the inclusion of new technologies different challenges arise. One example is the design of network interfaces (NIs) that could support and process packets up to 40-100 Gbps after the IEEE P802.3ba standard was released [31].

PAPER 3 tackles two additional research questions: 1) what is the impact of 100 Gbps on the network overall power consumption, and 2) what is the long term cost of using 100 Gbps technology. In order to evaluate the impact of using gross-grained technology on the network overall power consumption, the model based on the power-aware MILP formulation is used to determine which technology is the most power efficient for traffic demands from 2014 to a forecasted traffic demand in 2020. The forecasted traffic demand is based on Cisco's predictions [29]. Other MILP formulations to determined the most energy efficient configuration for MLR networks are found in literature [28] [32–35]. However, MILP formulations give the mathematical optimal solution without determining which is the gap between the best (i.e., the lowest power consumption) and the second best configuration. We push the state-of-art forward by running the model using a set of design constraints when defining the most power-efficient configurations to define which is the gross-grained designs impact on the network overall power consumption.

The power-aware MILP formulation is run three times. First, allowing the use of any technology (i.e., 10, 40, and 100 Gbps) defined as configuration 1 throughout this thesis, then allowing only the implementation of gross-grained devices (i.e., 40 and 100 Gbps) defined as configuration 2, and finally defining a whole infrastructure based exclusively on 100 Gbps technology defined as configuration 3. Note that the MILP formulation performs both network design and traffic engineering, which means that, it defines both which equipment has to be deployed in each node to satisfy all traffic demands, and which route to use for it. Routing is optimized to minimize power consumption, not using the standard shortest path algorithm. Next, once the equipment to be deployed in each node is defined, power consumption analysis for each network configuration from year 2014 to year 2020 is assessed.

Results show that the most energy-efficient technology is 10 Gbps due to its low granularity and efficient utilization of the slots from router chassis (which is the most power demanding element from the modular configuration described in Section 1.3.3). When using gross-grained designed (i.e.,

configuration 2), a power consumption increase between 2.6% and 8.8% is experienced. A network configuration solely using 100 Gbps technology turns to be non power-efficient, experiencing up to 32.4% power increase. Despite being the most energy-efficient technology (i.e., Joules/Watt) [36], 100 Gbps is underutilized with the short-term forecasted traffic demands, thereby leading to a higher overall power consumption.

In order to evaluate the long term cost of gross-grained designs the average nominal optical capacity per node (i.e., available capacity for the deployed equipment) is used as a figure of merit. Since only a 2-5% increase in the power consumption for the first five years is experienced when using configuration 2, it is of relevance to analyze how much additional capacity per node can be obtained with this configuration. Assuming that the average nominal optical capacity per node obtained by the configuration 1 to be the reference value, an average of 15.3% increase is obtained by using a gross-grained design. This additional nominal optical capacity provides 6 months to one year margin before the network needs to be upgraded (i.e., invest on new modules for the two-layer modular configuration, thus increasing CAPEX). Bigger margin, 2-3 years can be achieved by provisioning ahead using configuration 3, which provides 37.7% increase in the available optical capacity despite the additional power consumption expense. Based on the CAPEX values presented in [37], this can represent up to 25% CAPEX savings assuming no variations in the technology price.

In conclusion, although there are several papers that focused on minimizing power consumption in different layers, they were based on approximations on power consumption per port [38], or considered a single interface module [39]. **PAPERS 2-3** report a comprehensive CAPEX/OPEX analysis accounting for realistic and feasible configurations and evaluate the potential of high capacity devices (i.e., 40 and 100 Gbps technologies).

3.3 Transmission link power consumption

In WDM networks for medium and long transmission reach amplification is used. Thus, erbium-doped fiber amplifier (EDFA) is an essential component to conceive next generation high capacity WDM networks. Within the optical fiber transmission link EDFA contribution accounts for 14% of the overall power consumption [36], thus its study and characterization is of relevance . EDFA generalized models and models from a thermodynamic

point of view have been reported in [40] and [41], respectively. All reported models account only for a static power consumption value for EDFAs. In contrast, **PAPER 4** presents a model which accounts for both static and dynamic contributions, and evaluates the impact of the number of simultaneously amplified channels on the overall power consumption.

The analysis demonstrates that EDFA power consumption is dependent on the channel loading, which accounts for 48% of the overall power consumption. Carefully adjusting the pump laser power leads to a 67% reduction of the dynamic power consumption, thereby increasing global power savings.

3.4 KeyFlow for SDN in core networks

Network design trends focus currently not only towards cost-efficient solutions, but also towards strategies that provide the tools to cope with services flexibility requirements. Software defined networking is a programmable networking platform which enables an extra level of flexibility by detaching the control plan from the data plane. However, SDN for core networks presents a number of challenges since extending its coverage to large foot print networks arise concerns on scalability, QoS, and complexity [42] and [43].

Scalability is a well-known drawback for centralized control strategies such as generalized multi-protocol label switching (GMPLS) / path computation element (PCE) [44], since the centralized controller's complexity scales with the number of nodes . For the most used south-bound protocol within SDN, the OpenFlow [45], this scalability concerns are not only present at the centralized controller but also in the core forwarding engines, since the number of states related to active flows (i.e., flow tables' size) increases as the network scales up. This leads to a large number of lookup table operations for high node-density networks. Large number of operations translates into high node complexity, with the need of pricey and power demanding memories such as static random access memories (SRAMs), and ternary addressable content memories (TCAMs) [45; 46].

PAPER 5 and **PAPER 6** present a detailed analysis of an alternative protocol within SDN architectures named KeyFlow [15] from QoS perspective in terms of latency and jitter, and power consumption. KeyFlow enables the design of stateless low-latency router architectures where the

lookup process is replaced by a simple operation eliminating the need of SRAMs and TCAMs. This work promises a powerful approach for implementing SDN in core networks.

3.4.1 KeyFlow: operating principle

In this new south-bound protocol the aim is to reduce the core router's complexity (i.e., forwarding plane's complexity). In the standard OpenFlow protocol, core network nodes communicate with the OpenFlow controller when a packet with no associated matching rule reaches the forwarding engine. The controller sends the output port towards the next forwarding engine based on a defined routing algorithm, and this information is stored in the flow table from the core network node. For future flows with the same matching, the core node does not interact with the controller, instead checks the flow table which output port that packets needs to be forward to. When implementing KeyFlow in the core network forwarding engines, the traditional flow table lookup is replaced by a simple operation: a division.

As shown in Fig. 3.3, which is illustrative example of the KeyFlow operating system, the core network nodes (i.e., 3, 4, 5 and 7) do not communicate with the controller. Only edge nodes (i.e., ingress and egress nodes) do. The ingress node, node 1 in this specific example, interacts with the controller whenever there is non-matching rule for a specific packet that reaches the node. The KeyFlow controller computes a unique path ID, so-named *key*, based on the route that the packets has to follow to reach its

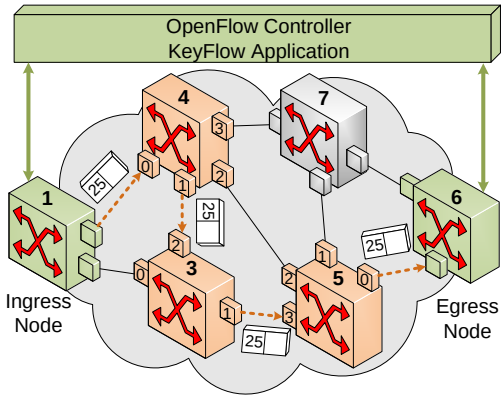


Figure 3.3: KeyFlow network architecture example.

destination (node 6 in this specific example). This algorithm has previously been referred as segment routing [47]. The ingress node incorporates a new rule in its flow table, which dictates that future packets with destination node 6 are assigned with the already defined path ID, i.e., key (25 in this specific example). The rest of core network nodes which have KeyFlow enabled do not have a flow table to lookup when a packet reaches them. Instead, they perform a division between the key value which is embedded in the packet header, and their own node ID. The reminder of the division determines the output port. Based on the example from Fig. 3.3, assuming that a packet from node 1 needs to reach each destination, node 6, through the intermediate nodes 4, 3, and 5, node 1 assigns key 25 to the packet, node 4 performs the operation 25 divided by 4. The reminder is 1, which is the output port towards node 3. Node 3 divides 25 by 3, with reminder 1, which is the output port towards node 5, and node 5 performs the division 25 by 5 which would result in the output port 0.

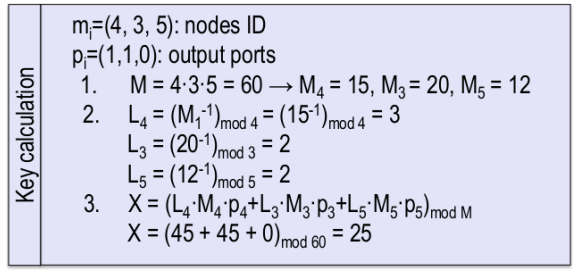


Figure 3.4: KeyFlow key calculation example.

As it is possible to extrapolate from the above example, the key assigned to each packet header is not a random number. It is a unique computed number based on the intermediate nodes' ID and their respective output ports. Let $< X >_i$ be defined as the remainder from X/i. In the KeyFlow architecture, X is the key value, and i represents each node ID. From the aforementioned example, the first intermediate node, node 4, performs the following operations $< 25 >_4 = 1$, which indicates that the output port is 1. Nodes 3 and 5 compute $< 25 >_3 = 1$ and $< 25 >_5 = 0$, respectively. For the key calculation, three steps are required. Fig. 3.4 summarizes the three steps, where $m_i = (4, 3, 5)$ are the nodes IDs, 25 is the key, and $p_i = (1, 1, 0)$ are the respective output ports. The first step consists on obtaining a value M_i for each node by dividing M by each node ID, where M is the product of all nodes IDs so that $M_i = M/m_i$. For example, for node 4, $M_4 = M/m_4 \rightarrow M_4 = 60/4 = 15$. In the second step, the parameter L_i

is calculated. L_i is the multiplicative inverse of M_i (i.e. $M_i \cdot L_i \equiv 1 \ (mod \ m_i)$). Finally, the value for the key, X is obtained by summing up the product of each L_i, M_i, and output port for each node, $L_i \cdot M_i \cdot p_i \ (mod \ M)$ [15].

3.4.2 Results and discussion

KeyFlow switch has been evaluated and compared with a reference OpenFlow switch. Table 3.1 summarizes the contributions reported in **PAPER 5** and **PAPER 6**, as well as the methodology and the platform used for each of the analysis. The evaluation includes latency, jitter, and power consumption results.

Table 3.1: KeyFlow contributions for SDN core forwarding engines compared to OpenFlow. PC, power consumption; RTT, round trip time.

	Methodology	Platform	Achievements
Latency	Emulation	Mininet	$>50\%$ RTT reduction for more than 3 hops
Jitter	Experimental	NetFPGA	Elimination of jitter for core nodes
PC	Experimental	XPower Analyzer (XPA)	57.3% power consumption reduction

Latency has been evaluated based on a round trip time (RTT) for a sent packet. Results are obtained based on an emulation platform called Mininet. RTT values for the reference OpenFlow switch are highly dependent on the number of hops, as well as the flow table density. Similar results are obtained for both OpenFlow and KeyFlow switches when the flow table of the OpenFlow switch is populated below 25%. For more populated flow tables (i.e., 50% and 75%) and more than 3 hops, more than 50% RTT reduction is observed when using the KeyFlow switch.

For highly populate flow tables, OpenFlow presents jitter values above 30 ms. This is an expected result since the RTT for a given packet depends on where in the table the matching rule is located. This jitter is completely eliminated when implementing KeyFlow, since no table lookup is performed, thereby the RTT for a given packet is always proportional

to the route length, number of hops, and the latency added by the key operation.

Finally, from a power consumption point of view, power savings of 57.3% are achieved with KeyFlow compared with the OpenFlow switch. This result is expected because with this simple forwarding plane, the need of power demanding TCAMs is eliminated. The TCAM is used for what is known as wildcard match which contributed with 78.75% of overall switch power consumption.

Chapter 4

100 Gbps IM/DD links for datacenters

The efforts in the design of short-range links for datacenters have been focused towards cost-effective solutions to provide high capacity and small foot-print solutions. In [17] an alternative solution for short-range high capacity links based on incoherent direct detection links using polarization multiplexed transmission is proposed. In this thesis this polarization-multiplexed solution is referred as quad-polarization. The contribution to the State-of-the-art includes a comparison between quad-polarization and two competitive technologies: parallel optics, and WDM. The conducted analysis includes computational complexity, power consumption, and performance in terms of receiver sensitivity. This chapter provides an overview on intensity modulated/direct detection (IM/DD) links used in datacenter scenarios, a description of the quad-polarization concept as well as requirements at both transmitter and receiver, and a summary of the main results obtained and presented in **PAPER 7**.

4.1 IM/DD links for datacenters

There is a need for high capacity links to cope with high bitrates in short-range links where coherent implementations are not attractive since size, cost, and power consumption of today's coherent transceivers are not suitable for datacenter networks. Research goes towards IM/DD systems to achieve low cost and power consumption solutions.

One approach to handle large volumes of data in short-range links is to increase the optical matrix size using parallel optical interconnects [48]. However, when dealing with datacenter environments, high density and small foot-print solutions are desired. An enabling technology for high capacity and density is WDM.

On one hand, WDM reduces the cabling overhead compared to parallel optics. On the other hand, current traffic demands challenge the achievable bit rate by WDM architectures. Consequently, different solutions to increase bitrates have been proposed, including the use of higher modulation formats, or space division multiplexing. Multi-level modulation formats such as 4 levels pulse amplitude modulation (4-PAM) [49], or duobinary modulation [50], have been explored to increase spectral efficiency. Research on more complex modulations formats such as discrete multitone (DMT) [51] or carrierless amplitude phase (CAP) [52] modulations have also been investigated.

Alternative to complex modulation formats, which come at the cost of computational complexity and power consumption [50], polarization multiplexed solutions have been proposed for IM/DD systems. Self-coherent direct detection solutions of an orthogonal frequency division multiplexed (OFDM) signal have been presented in [53]. Incoherent direction links based on the Stokes receiver reported in [54] and [55] were proposed to reduce complexity compared to self-coherent approaches [56].

Quad-polarization is an extension for incoherent direct detection links where four different data streams transmission is achieved by using four different states of polarization [17].

This thesis reports on the computational complexity, power dissipation, and performance in terms of receiver sensitivity of the quad-polarization scenario compared to the aforementioned parallel optics and WDM alternatives.

4.2 Quad-polarization concept

Fig. 4.1 illustrates the general concept of a quad-polarization system. The main idea is that four different data streams are transmitted simultaneously using the same media (in this case 2 km of standard single mode fiber (SSMF), using four different states of polarization: X-polarization (red), Y-polarization (blue), 45°-polarization (yellow), and left circular (LC) (green).

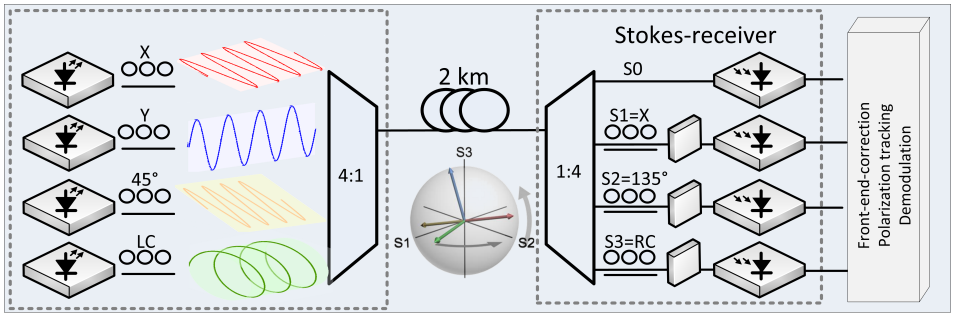


Figure 4.1: Conceptual illustration of the quad-polarization system for IM/DD optical data links.

At the receiver the X, the 135°, and the right circular (RC) states of polarization are detected (or the orthogonal counterparts) together with the instantaneous total intensity. The Stokes analyzer is used for detection. Next, in order to demultiplex the four states of polarization two key-steps are performed: 1) compensating for the rotation that the states of polarization experience when transmitted over SSMF (as indicated by the vectors in the Poincaré sphere in Fig. 4.1), and 2) demodulation in the digital domain using low-complexity DSP.

4.2.1 Stokes analyzer

The Stokes analyzer is composed of four branches designated as S0, S1, S2, and S3. At the first branch, S0, the instantaneous total intensity is measured. The other three branches measure the optical power based on Equations 4.1, 4.2, and 4.3, respectively

$$S1 = 2I_{X|Y} - S0 \tag{4.1}$$

$$S2 = 2I_{45|135} - S0 \tag{4.2}$$

$$S3 = 2I_{RC|LC} - S0 \tag{4.3}$$

where S1, S2, and S3 measure the optical power associated to X or Y-polarizations, 45°or 135°, and right- or left-circular, respectively. The full operation in Stokes space along with the inclusion of circular state of polarization (green state of polarization in Fig. 4.1) enabled both the addition of a fourth state of polarization endowing the system with an extra

channel, and polarization rotation compensation without the need for the phase information, thus obtaining a non-coherent solution for short-range optical transmission systems.

4.3 Digital signal processing

PAPER 7 presents both complexity and power dissipation analysis for parallel optics, WDM, and quad-polarization based on the DSP presented in this Section. The evaluation is assessed by breaking down each of the DSP blocks in the number of real additions and multiplications required.

Three different stages are required in the DSP to successfully demodulate the data from the quad-polarization system. Fig. 4.2 provides a general overview of the three stages which enclose general front-end correction, polarization tracking, and demodulation. General front-end corrections are common DSP blocks for all the analyzed scenarios: WDM, parallel optics, and quad-polarization. It includes error estimation (i.e., timing error correction) and resampling to the minimum number of integer samples per symbol.

Specific for the quad-polarization system, an additional stage is required for polarization tracking which is used to compensate for the Poincaré sphere rotation after fiber propagation. The used polarization tracking algorithm is based on the work presented in [56]. While the algorithm described in [56] considers only the two orthogonal states of polarization, in [17] an extension is done enabling demodulation of four different states of polarization. A more detailed block diagram of the second stage for polar-

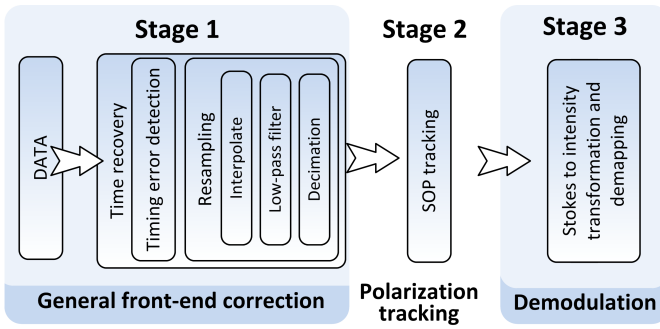


Figure 4.2: DSP blocksfor 4-SOP IM/DD. DSP, digital signal processing; IM/DD, intensity modulated/direct detection; SOP, state of polarization.

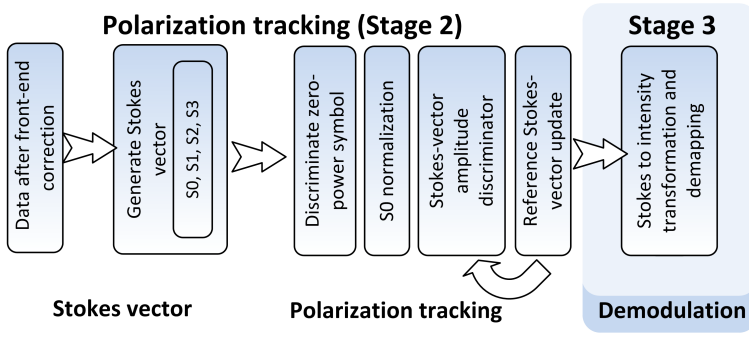


Figure 4.3: DSP blocks for stage 2: state of polarization tracking

ization tracking is presented in Fig. 4.3. First, the Stokes vectors are generated based on Equations 4.1, 4.2, and 4.3. Next, all symbols with intensity below given threshold S_{th} are considered zero-power symbols, thereby demapped and removed from the Stokes sequences to avoid mathematical indeterminates during the normalization process. S0-normalization follows to maximize the polarized components power. Finally, an iterative process is performed involving a Stokes-vector amplitude discriminator and a reference Stokes-vector update. The objective of these two sub-blocks is to enable tracking of the Stokes vectors associated with two or three transmitted signals.

Finally, the third stage of the DSP includes Stokes to intensity transformation and de-mapping. Demodulation is based on a 4x4 multiple-input-multiple-output (MIMO) process. The computational load of the combined de-mapping and derotation is low since the mapping from the transmitter states of polarization to the Stokes apace and its inverse are known a priori. Additionally, digital processor speed and power consumption requirements are mitigated since no need for real time matrix inversion is needed because the derotation matrix is orthogonal.

4.4 Results and discussion

Quad-polarization system has been evaluated and compared with two reference systems: parallel optics and WDM. Table 4.1 summarizes the contributions reported in **PAPER 7** as well as the methodology used for each of the analysis assessed. Evaluation encloses computational complexity, power dissipation, and receiver sensitivity results.

Table 4.1: Quad-polarization evaluation as compared to parallel optics and WDM systems. ADC, analog to digital converter; DSP, digital signal processing; WDM, wavelength division multiplexing.

	Methodology	Achievements
Receiver Sensitivity	Experimental	-10.5 dBm for parallel optics -9.5 dBm for WDM 4.4 dBm for quad-polarization
Computational complexity	Numerical analysis	>1% computational complexity increase associated to the DSP
Power consumption	Numerical analysis	47.6% power consumption increase due to ADCs

For the receiver sensitivity, an experimental evaluation for three different testbed has been conducted. The configuration analysis has been assessed for 4x32 Gbps back-to-back (B2B) and for 2 km transmission of standard single mode fiber type G.652. The receiver sensitivity reported here accounts for the entire system performance rather than the response for each individual photodiode (variations in sensitivity between channels is observed for all systems).

Computational complexity associated to the additional DSP needed for state of polarization tracking and demodulation in the quad-polarization system is negligible compared to the front-end-correction. Its contribution is below 1% of the overall number of real operations. However, a significant power consumption increase is experienced due to the power demanding analog to digital converters required for demodulation in the digital domain.

Quad-polarization represents an important advance within the telecommunication research area. It removes barriers to progress since combining quad-polarization with WDM, and/or parallel optics low complexity 400 Gbps solutions can be achieved for short-reach optical transmission systems. The contribution presented in this thesis includes the evaluation of its power consumption, computational complexity, and receiver sensitivity. It shows how quad-polarization enables bitrate quadrupling for each data lane yet at power consumption and performance costs. The performance degradation observed is typically associated with multilevel modulation formats.

Chapter 5

Conclusion and future work

5.1 Conclusions

This thesis is based on a multidisciplinary approach which provides a extensive understanding of power consumption in different network segments. The research results presented in this thesis are pioneering providing network designers with two different strategies for implementing energy efficient solutions. First, techniques energy-efficient solutions that exploit already existing technologies or current deployments. Second, modeling for the network of the future providing evaluation performance and power consumption of breakthrough technologies such as software defined networking (SDN) for core networks or quad-polarization for short-range optical links.

These results have a scientific and technological impact since due to the resources scarcity and the energy consumption regulations trends, power awareness is thriving in optical network design strategies. Thus, the scientific novelty of sliceable transceivers, KeyFlow, and quad-polarization is not limited to cost savings, latency reduction, or bandwidth enhancement, it also provides technical insights for energy efficiency. In addition, this work has a strategic and economic impact, since the reported results have the potential to boost the industrial production of products based on power-aware technologies.

One of the main strongholds of this thesis is that it accounts, throughout the entire research approach, for fundamental trade-offs such as power consumption, latency, complexity, and cost.

5.1.1 Re-use of current technologies for energy-efficiency

PAPERS 1, 2, and **4** offer solutions to conceive design by re-using already deployed equipment, defining a network infrastructure and traffic engineering with mature and well-established technologies, and providing scientific understanding from a power consumption point of view of fundamental optical network components. The advantage of incorporating power-awareness in various areas where network providers have already a rooted and firm expertise facilitates the actual implementation of the proposed solutions.

The sliceable transceiver introduced in **PAPER 1** exploits the benefits of re-using low bandwidth equipment with natural reallocation on demand to satisfy current traffic demands. The breakthrough concept is to provide a long run action plan which focuses on cost savings when upgrading the network. These savings are achieved accounting for an expected cost reduction as new technologies (e.g., 100 Gbps technology) move from the first-generation product, at a high price and where lots of customization is needed to a more mature stage.

PAPER 2 presents a new model for network design based on a traditional well-established WDM technology. An integrated analysis from CAPEX and OPEX perspective is discussed and trade offs such as cost, network reconfiguration, and power consumption are addressed.

Finally, in **PAPER 4**, a comprehensive understanding of the fundamental optical amplifier, EDFA, from a power consumption perspective is reported. A power consumption analysis based on empirical data is presented together with a mathematical model. This contribution represents a key-tool for network operators to better utilized and manage already deployed network resources.

These three contributions to the state-of-the-art account for two relevant trade offs: first, power consumption in contrast to quality of service in terms of latency, and second, power consumption and cost. The outcome is a range of solutions for network providers to deliver different class of services.

5.1.2 Inclusion of new cutting-edge technologies

The main idea of the second set of results presented in this thesis relies on the concept that current network infrastructures are not capable to support future bandwidth capacity and flexibility requirements, as well as easy manageability of the network infrastructure. Consequently, new technologies and boost of arising breakthroughs need to be evaluated before envision their inclusion in the Internet.

PAPER 3 encloses a more concrete target, 100 Gbps technology which has the potential to accommodate future traffic demands. The model from **PAPER 2** is expanded considering gross-grained network designs where 40 and 100 Gbps equipment is used exclusively. Results show that additional 37.7% optical network capacity is achieved when using solely 100 Gbps technology at power consumption expenses between 13.3% and 32.4%. The long term vision of high capacity equipment is CAPEX savings associated to the additional optical bandwidth provided by gross-grained designs, where up to 25% CAPEX savings are obtained by provisioning up to 3 years ahead. This analysis highlights an interesting trade-off between cost and power consumption, since the most cost-effective solution is not always the least energy consuming.

Along the lines of promising emerging technologies, SDN is considered as a potential alternative solution for control of transport networks in **PAPER 5** and **PAPER 6**. This architectural proposition aims at enhancing flexibility in today's control plane to cope with current applications on-demand requirements. The analyzed south-bound protocol for SDN named KeyFlow satisfies the specific challenge of reducing forwarding fabric engine complexity, latency, jitter, and power consumption compared to the reference protocol OpenFlow. KeyFlow represents a *key enabling technology (KET)* for SDN in transport networks.

Finally, **PAPER 7** evaluates a novel polarization multiplexing strategy for high capacity short range optical links compared to two alternative solutions: parallel optics and WDM. Quad-polarization had been presented as a proof-of-principle showing the potential of transmitting four independent data streams through a single media. The contribution summarized in this thesis includes a comprehensive analysis from a complexity, power consumption, and performance point of view. When a technology trigger arise, a comparison with competitor technologies is required so that potential audience moves towards its adoption. While quad-polarization

goes beyond the state-of-the-art short range optical transmission links, the work described in this dissertation shows its frailty, including high receiver sensitivity and increased power consumption due to the power-demanding ADCs. On the other hand, the complexity added for polarization tracking accounts for only 2.47% of the total power consumption, showing the potential for quad-polarization implementation in current infrastructures.

5.2 Future work

Research on energy-efficiency in optical networks has been a research line since early 2000s. However, it is only recently that telecommunication operators and Internet service providers have considered energy-efficiency for wired networks and service infrastructures a relevant target.

5.2.1 Short-term vision

Already existing telecommunication operators need to be encourage by the potential OPEX savings, yet GNTs have to consider the risk of the investment plan. The first part of this thesis which was focused on strategies that consider current deployed infrastructures and already established technologies can be exploited further. When it comes to re-using low bandwidth equipment by means of sliceable transceivers the work presented was solely based on NRZ-OOK modulation formats. Further power consumption reduction can be achieved by minimizing the energy to process a bit by exploring higher order modulation formats such as duo-binary, M-PAM, or CAP modulation formats.

Regarding MLR configurations, in this thesis resource consolidation (i.e., grooming in translucent networks to reduce power consumption) and adaptive link rate to accommodate dynamic traffic demands has been presented. For this, three thresholds are defined: low, average, and peak traffic loads. This work can be extended by defining a real-time heuristic. MILP models provide the optimum configuration, but they have slow convergence time due to their computational complexity. Hybrid solutions which benefit from the introduced MILP formulation and simpler heuristics could boost potential power savings in well-established WDM networks based on real-time adaptation. Flexible networking in conjunction with flexible transceivers are an appealing alternative to MLR configurations to be considered.

5.2.2 Long-term vision

Longer term optical networks are envisioned as intelligent systems capable to self-reconfigure and self-manage resources to cope with current traffic demands in a power-aware fashion. Two key concepts are aligned to smart-networks: SDN, which has been tackled in this thesis, and cognitive optical networks. Software defined cognitive optical networks (SD-CON) benefit from the flexibility provided by software both at the transport and control planes and the intelligence given by cognitive techniques which are based on human-brain abilities integrated into optical networks. Cognition consolidated with SDN will allow for a great overview of the network resources and their utilization, enabling optimum integrated data-signaling and control framework which can support emerging services adaptively.

SD-CON are expected as a breakthrough technology for future optical networks. KeyFlow, analyzed in this thesis, is the enabling south-bound protocol for SDN in transport networks, but there is still an important need to develop cognitive equipment to be incorporated in current optical network infrastructures. There are other relevant research challenges to be address, such as reducing complexity of the SDN controller and map intelligence in SDN forwarding engines to exploit the full potential of cognitive optical networks.

Paper 1: Digital Signal Processing for a Sliceable Transceiver for Optical Access Networks

Silvia Saldaña Cercós, Christoph Wagner, Juan José Vegas Olmos, Anna Manolova Fagertun, and Idelfonso Tafur Monroy, "Digital Signal Processing for a Sliceable Transceiver for Optical Access Networks," *ISCC*, 2015.

Digital Signal Processing for a Sliceable Transceiver for Optical Access Networks

Silvia Saldaña Cercós, Christoph Wagner, Juan José Vegas Olmos, Anna Manolova Fagertun, and Idelfonso Tafur Monroy
Department of Photonics Engineering, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark,
e-mail: (ssce@fotonik.dtu.dk)

***Abstract*—Methods to upgrade the network infrastructure to cope with current traffic demands has attracted increasing research efforts. A promising alternative is signal slicing. Signal slicing aims at re-using low bandwidth equipment to satisfy high bandwidth traffic demands. This technique has been used also for implementing full signal path symmetry in real-time oscilloscopes to provide performance and signal fidelity (i.e. lower noise and jitter). In this paper the key digital signal processing (DSP) subsystems required to achieve signal slicing are surveyed. It also presents, for the first time, a comprehensive DSP power consumption analysis for both WDM and TDM systems at 1 Gbps and 10 Gbps, discussing latency penalties for each approach. For 1 Gbps WDM system 278 pJ per information bit for 4 slices is reported at 105 ns latency penalties, whereas 3898.4 pJ per information bit at 183.5 μs latency penalty is reported for 10 Gbps. Power savings of the order of hundreds of Watts can be obtained when using signal slicing as an alternative to 10 Gbps implemented access networks.**

***Keywords*—digital signal processing; signal slicing; power consumption; access networks.**

I. Introduction

More than 40% of the world population are internet users [1] . The number of internet users has a growth rate of 8% per year due to the variety of internet access technologies as well as the emergence of new applications such as social networking, online gaming, video streaming, e-health systems, or video conference. The increasing number of users pushes further the traffic demands with a forecasted IP traffic growth of 21% for the next five years [2]. Consequently, network providers are forced to upgrade their networks. Commonly, network capacity upgrades are achieved by replacing low bandwidth electronics or optical equipment with high bandwidth technology. However, the network equipment costs follow an approximately linear relationship with the traffic demand [3], thus higher bandwidth equipment becomes a non cost-effective solution. Therefore, it is of interest to study capacity upgrading strategies that re-use to a large extend existing low bandwidth technology and yet satisfy the service requirements and costs constraints in an energy efficient manner.

In [4] an alternative capacity upgrade strategy based on sliceable transceivers is presented. The advantageous features of this strategy includes re-using low bandwidth equipment with natural bandwidth reallocation on demand to satisfy dynamic traffic demands. Moreover, electronic bandwidth is a major limiting factor in optical access networks which cause a bottleneck in optical access networks [5]. Using sliceable transceivers avoids network bottlenecks while accommodating all demands at the apex of traffic loads. This strategy overcomes the disadvantages of over-provisioning the network to cope with traffic demands such as under-utilization of network resources for low traffic hours.

Besides arising the need of upgrading strategies, traffic growth leads to a corresponding increase in power consumption which triggers concerns on electricity generation limitations. Even though for the past 10 years there has been an electricity generation growth of 3% [6], since power consumption increases proportionally to the traffic growth, it is expected that by year 2022 network equipment power consumption will reach the total world electricity generation [6]. Consequently, before adopting any new solution for coping with the current traffic demands, it is important to evaluate its impact on the overall network power consumption. Thus, power consumption has become a major network design factor.

Even though higher bandwidth technology is more power efficient (i.e. W/Gbps) [7], it can represent a power expense at low traffic-loads (e.g., during the night, weekend, or vacation periods) when the equipment is under-utilized. On one hand, lower bandwidth devices enabled by the use of sliceable transceivers could strengthen the benefits of on/off sleep mode strategies, thereby reducing the overall network power consumption. On the other hand, signal slicing comes at the expenses of digital signal processing (DSP) both at the transmitter and the receiver. This DSP algorithm has been also used in a real time oscilloscopes to provide with full signal path symmetry which offers lower noise and jitter, and higher bandwidth (i.e. 70 GHz) [8].

This paper presents a detailed description of the required DSP for sliceable transceiver for optical access networks as well as an analysis of the power consumption associated to the sliceable transceiver DSP for both 1 Gbps (experimental-based) and 10 Gbps (simulation-based). The DSP power consumption analysis is based on the needed number of operations. This paper also highlights where energy savings can be achieved at which latency penalties defining future areas for research.

II. Signal slicing digital signal processing

This section provides the needed background on the signal slicing digital signal processing for its power consumption evaluation. It offers the overall system overview as well a description of each of the DSP blocks required to implement sliceable transceivers.

Signal slicing breaks a high bandwidth signal into a set of lower bandwidth signals. The creation of low bandwidth slices

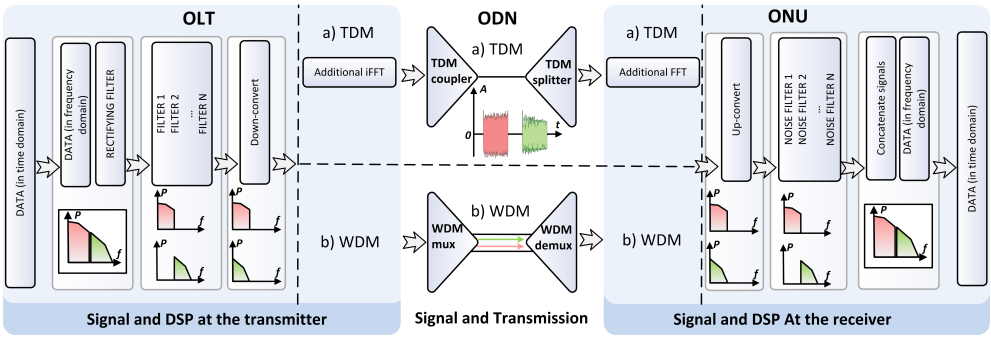


Fig. 1: DSP in signal slicing. DSP, digital signal processing; FFT, fast fourier transform; iFFT, inverse fast fourier transform; OLT, optical line terminal; ODN, optical distribution network; ONU, optical network unit; TDM, time division multiplexing; WDM, wavelength division multiplexing

from a wide bandwidth signal, and the reconstruction process at the receiver side requires the use of DSP. Fig. 1 presents the DSP needed in signal slicing for downlink in optical access networks from the transmitter (i.e. optical line terminal (OLT)) to the receiver (i.e. optical network unit (ONU)). For the optical distribution network (ODN), two different approaches are represented for comparison purposes. The first approach for the ODN is based on the legacy time division multiplexing (TDM) optical access network depicted in Fig. 1 a), and the second strategy based on wavelength division multiplexing (WDM) shown in Fig. 1 b).

As illustrated in Fig. 1 within the inner transmitter there are several blocks or subsystems which are dedicated to the signal slicing process. First, filtering is implemented in the frequency domain, thus before starting with the actual signal processing, a fast Fourier transform (FFT) is applied to the data in the time domain. A rectifying filter is needed in this process in order to cut off the mirrored sideband due to the FFT. Then, a set of *n* filters are implemented in order to achieve *n* slices from the wide bandwidth signal. In order to benefit from low bandwidth technology before transmitting each slice, each of the high frequency slices is down-converted to base-band. The transmission of the signal slices can be done in a number of ways. In the first case, using legacy TDM channels, an additional inverse FFT (iFFT) is needed in order to return the signal to the time domain. Experimental signal slicing and transmission in time domain has been reported in [4]. However, TDM systems are time sensitive and require synchronization and scheduling techniques so that each slice subsists in the channel for only a fraction of time. An alternative solution to TDM is using different wavelengths for each signal slice which enables parallel transmission using different optical carriers. WDM optical access networks boost upgrading TDM optical access networks to next generation access networks (NGA) where an increase in capacity and reach is achieved. In this second case no additional iFFT is required prior to the WDM multiplexer.

At the inner receiver the transmitter complementary blocks or subsystems are required. First, each base-band slice needs to be up-converted to its original frequency band. Subsequently, a set of *n* filters are implemented in order to suppress the noise added during the signal transmission. Finally, the signal slices are concatenated and transformed back into the time domain.

A. Power consumption analysis of the DSP

One of the key issues for adopting sliceable transceivers as an upgrading strategy is the additional power consumption required for the DSP needed for the signal slicing. The DSP power consumption analysis is evaluated by calculating the required number of integer operations performed by each subsystem at the transmitter and at the receiver. The number of needed integer operations (i.e., real additions and real multiplications) depends on the specific implementation of each subsystem. The results obtained in this paper are based on the proof-of-concept presented in [4] where no power consumption optimization has been studied. Additionally, the power consumption associated to the DSP is also dependent on the specific application-specific integrated circuit (ASIC). Consequently, the method outlined herein is designed to provide an overview of each subsystem power consumption contribution rather than give exact estimates. The energy consumption of the DSP E_{DSP} is given by equation 1.

$$E_{DSP} = E_{FFT}+E_{rf}+E_{dc}+E_{uc}+E_{nf}+E_{con}+E_{iFFT} \quad (1)$$

where E_{FFT}, E_{rf}, E_{dc}, E_{uc}, E_{nf}, E_{con}, and E_{iFFT} are the energy consumption for the FFT, rectifying filters, down-conversion, up-conversion, noise filters, concatenation, and the iFFT, respectively.

Let E_m and E_a be the energy in Joules that each real multiplication and real addition consumes, respectively. The total energy consumed by each DSP subsystem E_i, where $i \in [FFT, rf, dc, uc, nf, con, iFFT]$, is calculated as represented in equation 2, where Q_i stands for the number of real multiplications and A_i for the number of real additions required by the DSP subsystem E_i.

$$E_i = Q_i E_m + A_i E_a \quad (2)$$

In this section, a description of each DSP subsystem in terms of number of integer operations is provided. The first subsystem consists on an FFT to transform the signal from the time domain into the frequency domain. The number

of operations required in order to perform the FFT depends on how the FFT is implemented, e.g., whether it is radix-2, radix-4, or split-radix [10]. Additionally, the scheme used to perform complex operations has also an impact on the overall number of operations. In this paper, a split-radix FFT using a three butterfly structure with a 4-multiply 2-add scheme per complex multiplications is assumed [11]. The presented analysis provides a lower boundary, since radix-2 or radix-4 FFT implementations would increase the number of integer operations, thereby power consumption. Equations 3 and 4 can be used to generate counts for real multiplications and real additions, respectively [11].

$$Q_{FFT} = (4/3)MN - (38/9)N + 6 + (-1)^M(2/9) \quad (3)$$

$$A_{FFT} = (8/3)MN - (16/9)N + 2 + (-1)^M(2/9) \quad (4)$$

where *N* stands for the length of the FFT, and *M* for the number of stages, so that $N = 2^M$.

The rectifying filter energy consumption contribution depends on its number of complex taps used to cut off the mirrored sideband due to the FFT. Assuming *N* number of complex taps and using four real multiplications and two real additions per complex multiplication, equation 5 accounts for the energy consumption of n rectifying filters, where n is the number of slices obtained from the wide bandwidth signal.

$$E_{rf} = nN(4E_m + 2E_a) \quad (5)$$

The last block in the transmitter side performs down-conversion of each slice before transmission. This is achieved by circularly shifting the bits of each slice. Bits shifting is assumed to be a pointer change, thus not involving any addition or multiplication in the process. Shift operation involves register read and writes. The average energy consumption per register read and write E_{opR} is $3.43 \cdot p_t \cdot V^2$, where p_t is the CMOS process technology feature size in nm, and V is the CMOS supply voltage for a given process technology [12]. The energy consumption for down-converting E_{dc} is represented in equation 6 where n-1 slices of nb bits need to be down converted to baseband.

$$E_{dc} = E_{uc} = (n-1)nbE_{opR} \quad (6)$$

where nb is the total number of bits divided by the total number of slices.

In case of using legacy TDM in the ODN, an additional iFFT is required at the transmitter side. The number of operations of the iFFT in this case is based on equations 3 and 4 where *N* stand for the length of the FFT per slice.

At the receiver side the complementary subsystems are needed. Like the down-conversion subsystem, the up-conversion subsystem is based on a circular bit shifting, and its energy consumption contribution, E_{UC} is given by equation 6. The energy consumption of the n noise filters is also based on their number of complex taps. Thus, E_{nf} is based on the number of real multiplications and real additions presented in equation 5.

An additional subsystem is required at the receiver side which corresponds to the signal concatenation procedure. The number of additions required depends on the number of slices as presented in equation 7 where n stands for the number of slices. Equation 7 is only valid assuming that all the slices have the same length (s_{length}), and that there is a minimum of two slices.

$$A_{con} = s_{length} \cdot \sum_{i=2}^{n} i \quad (7)$$

To retrieve the signal in time domain, an iFFT is performed at the receiver side. Analogous to the FFT, the number of real additions and real multiplications depends on the FFT implementation. In this paper we asume FFT and iFFT implemented in the same way, thereby assuming the same number of real multiplications and real additions as the ones presented in equations 3 and 4, respectively.

III. Results

In this section the model developed in the previous section is used to analyze the DSP power consumption for sliceable transceivers. Two different cases are evaluated: 1) The use of sliceable transceivers for 1 Gbps non return to zero on off keying (NRZ-OOK) based on the experimental work presented in [4], and 2) the use of sliceable transceivers for 10 Gbps NRZ-OOK. This second scenario is simulation-based in order to evaluate the plausible benefits of sliceable transceivers as an alternative solution to 10 Gbps TDM PON networks.

We consider use of an ASIC based on 90 nm CMOS process technology. There is a commercially available ASIC designed by Nortel, which published a significant amount of detail including the energy required to process a bit of information, being 217 pJ per bit, and the energy required to perform a computation, being 0.83 pJ per integer operation. After doing some reverse engineering [9] reports 1.5 pJ and 0.5 pJ average energy consumed per real multiplication and real addition, respectively. The energy results presented in this section are based on these values. In this work, the DSP subsystems have not been optimized for power consumption. Additionally, the exact power consumption of the DSP depends on the specific ASIC implementation. Thus, reported results provide an estimate rather than exact computations.

A. Sliceable transceiver for 1 Gbps NRZ-OOK signal

The parameters used for modeling the energy consumption are the ones presented in [4] with 1 Gbps NRZ-OOK signal.

Fig. 2 plots the energy consumed per information bit as a function of the sampling frequency for the WDM (square, green), and TDM (circle, blue) systems. The sampling frequency has been swept from the Nyquist limit (i.e., 2 GHz) up to 10 times the signal bandwidth. This plot is based on equation 1 for 1 Gbps NRZ-OOK signal sliced into two slices. Results shows that WDM system consumes 218.45 pJ per information bit outperforming TDM by 69.54% which consumes 717.16 pJ per information bit at the Nyquist limit. Improvements are reduced to 51.83%, 33.72%, and 19.43% when oversampling the signal by 4, 6, and 10 GHz, respectively. Oversampling for improving the resolution, reducing the noise, reducing aliasing and relaxing the anti-aliasing filter requirements comes at a

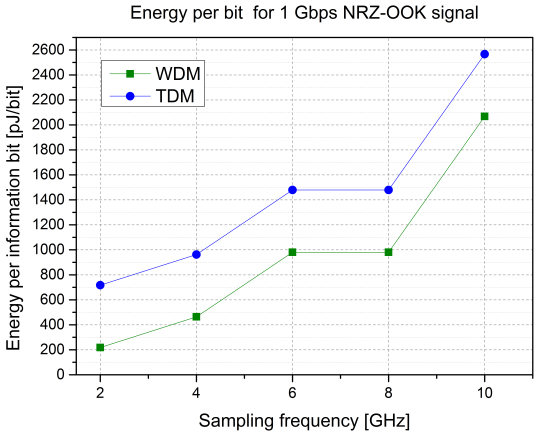


Fig. 2: Energy per information bit as a function of the sampling frequency for WDM and TDM systems for 1 Gbps NRZ-OOK signal.

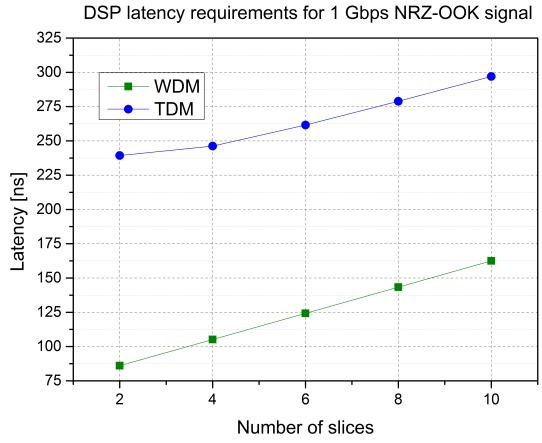


Fig. 4: DSP latency requirements as a function of the number of slices for WDM and TDM systems for 1 Gbps NRZ-OOK signal.

expenses of power consumption, obtaining more than double energy per bit information bit when oversampling at 4 GHz.

Increasing the number of slices helps relaxing the equipment bandwidth requirements. However, it also increases the computational complexity of the DSP leading to an increase in energy consumption. Fig. 3 illustrates the energy consumption per information bit as a function of the number of slices for both WDM (square, green), and TDM (circle, blue) systems. From the subsystems described in Section II, there is an increase on the number of integer operations for the rectifying and noise filters, and the concatenation process as presented in equations 5, and 7, respectively. Results show a 30 pJ increase per additional slice, obtaining 278 pJ per information bit for 4 slices for a WDM system. Bandwidth requirements relaxation also comes at performance expenses in terms of latency. In

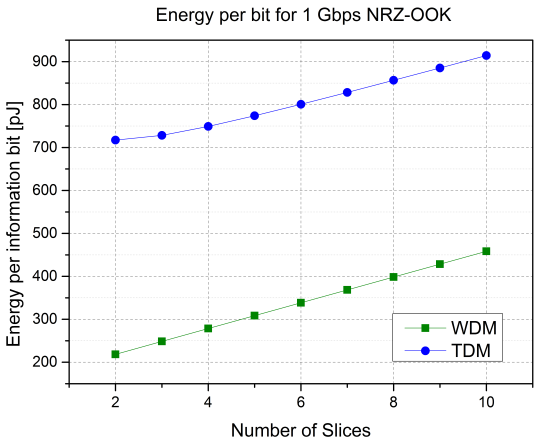


Fig. 3: Energy per information bit as a function of the number of slices for WDM and TDM systems for 1 Gbps NRZ-OOK signal.

order to evaluate the system DSP latency requirements it is assumed that the specific ASIC performs 12 trillion integer operations per second [9]. Accounting for the number of integer operations described in Section II-A, Fig. 4 is derived. The results presented in Fig. 4 only account for the DSP latency. As expected WDM outperforms TDM offering 65% lower latency. Results show 105 ns and 246 ns for 4 slices in WDM and TDM systems, respectively. In case of TDM systems additional latency needs to be accounted for the time that each data slice takes to reach the destination.

B. Sliceable transceiver for 10 Gbps NRZ-OOK signal

To show the impact of the proposed approach, the method is evaluated at higher bit rates (i.e., 10 Gbps) where bandwidth becomes a major limiting factor. Unlike sliceable transceivers at 1 Gbps which had been experimentally demonstrated [4], 10 Gbps sliceable transceivers results are based on simulations. This evaluation aims at offering an overview of signal slicing as an alternative to real implemented access networks such as 10 Gbps TDM PON. Energy per bit information and latency are evaluated for WDM and TDM systems for two different sampling frequencies: 20 GHz and 160 GHz. 20 GHz has been selected as it complies with the Nyquist criterion. However, after signal restoration at the receiver side, significant signal degradation is observed. Thus, results at 160 GHz sampling frequency are also presented where a clear and open eye after signal restoration was obtained.

Fig. 5 illustrates the energy per information bit as a function of the number of slices at 20 GHz (solid pattern) and 160 GHz (diagonal straps) sampling frequencies for WDM (square, green) and TDM (circle, blue) systems.

Similar energy consumption is obtained for both WDM and TDM systems. At 20 GHz sampling frequency the energy per bit is below 418 pJ per information bit for both WDM and TDM systems, both for 2 and 3 slices cases, and the maximum number of slices to ensure an open eye after restoration is 3. In order to increase the number of slices 160 GHz sampling frequency is required. Fig. 6 shows the eye diagrams of a

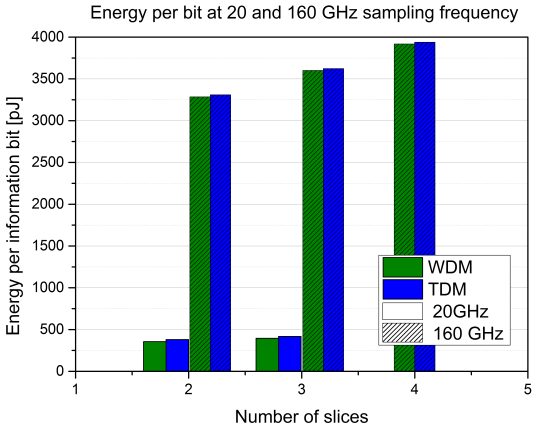


Fig. 5: Energy per information bit as a function of the sampling frequency for WDM and TDM systems for 10 Gbps NRZ-OOK signal.

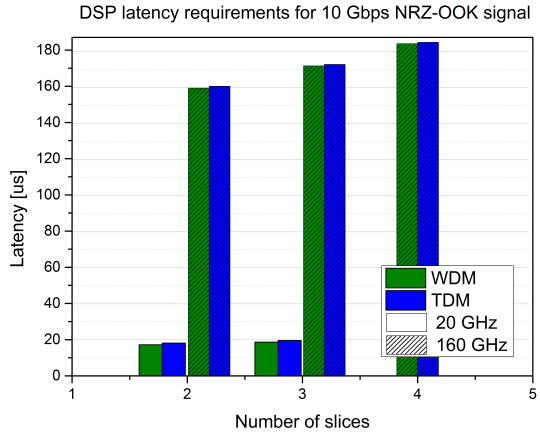
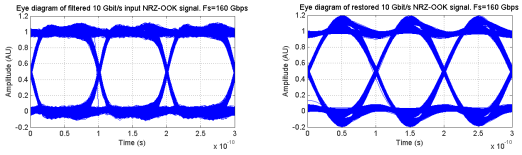


Fig. 7: DSP latency requirements as a function of the number of slices for WDM and TDM systems for 10 Gbps NRZ-OOK signal.

10 Gbps NRZ-OOK signal sampled at 160 GHz, where 6a) is the filter signal at the receiver side before signal slicing is applied, and 6b) is the eye diagram at the receiver side after restoration of 4 slices. Four slices are chosen since they allow the use of equipment for 2.5 Gbps transmission (used for Gigabit-PON (GPON) standards) to achieve bitrates up to 10 Gbps. The increase in the sampling frequency to ensure an eye opening at the receiver side comes at expenses of power consumption. As illustrated in Fig. 5, for 4 slices, 3898.4 pJ and 3919.1 pJ per information bit are consumed for WDM and TDM systems, respectively. These results are one order of magnitude higher than for the 1 Gbps case presented in Section III-A due to the increased computational complexity associated to the FFT, rectifying and noise filter described in Section II-A.

A performance comparison in terms of DSP latency is presented in Fig. 7. Latencies below 20 μs are obtained for 20 GHz sampling frequency for all cases. For 160 GHz and 4 slices (which would provide with the tools to implement 10 Gbps access rate with a 2.5 Gbps system, relaxing equipment bandwidth requirements by 4) 183.5 μs and 184.29 μs latency are obtained for WDM and TDM systems, respectively.



(a) (b)

Fig. 6: Eye diagrams of 6a) a filtered 10 Gbps NRZ-OOK signal at the receiver before signal slicing is applied and 6b) after restoration at the receiver side.

Latency penalties are significantly higher as the bit rate increases, obtaining more than three orders of magnitude higher latency for a 10 Gbps sliced signal compared to 1 Gbps sliced signal.

C. Complete system overview

The work presented in [4] shows one possible solution for the signal slicing implementation. Reporting results on the experimental demonstration of a scalable sliceable transceiver for intensity modulated optical access networks is out of the scope of this paper. However, in order to provide a complete power consumption comparison between high and sliced bandwidth signals it is important to consider the different electrical and optical components' contributions. The experimental setup reported in [4] includes, at the OLT, a binary data stream generator, and arbitrary waveform generator (AWG), and a light source. The data stream from the binary data generator is sent to a DSP block. The different sliced signals are sent to an AWG. Its output is amplified through an electrical amplifier. A distributed feedback laser -electroabsorption modulator (DFB-EAM) is used as an optical source. The modulated signal is launched into a single mode fiber (SMF). At the ONU side, direct detection is employed using a positive intrinsic negative photo diode (PIN PD). The electrical baseband signal is captured by a digital storage oscilloscope. Finally, offline DSP is applied in order to reconstruct the initial signal. Further details on the experimental demonstration can be found in [4].

Figs. 2, 3, and 5 show an increase in energy consumption due to the additional DSP needed for signal slicing. On the other hand, slicing the signal represents a mitigation on the optical an electrical equipment requirements. From all the electrical and optical components mentioned above the amplifier, the DFB-EAM, and PIN PD present a negligible variation in terms of overall power consumption when increasing bandwidth requirements. However, the AWG has a relevant impact. Table I lists several commercially available power consumption values for different bandwidth requirements for

TABLE I: Arbitrary waveform generator power consumption.

Model	Bandwidth	Power consumption	Reference
Agilent M8190A	5 GHz	210 W	[13]
Agilent 81180B	2 GHz	100 W	[14]
Agilent 33600A	120 MHz	75 W	[15]
Agilent 33500B	30 MHz	45 W	[16]

the AWG for indication purposes. Assuming a relaxation in the AWG bandwidth requirements from 120 MHz to 30 MHz, which would be achieved by slicing the signal in 4, represents 30 W reduced power consumption. For a 1 Gbps NRZ signal with 4 slices a power consumption penalty below 280 mW is presented in Fig. 3 indicating potential power savings by using sliceable transceivers at 1 Gbps. This power savings come at a low latency penalty, 105 ns for 4 slices in a WDM system. When it comes to higher bit rates scenarios AWG power savings when lowering the bandwidth requirements are hundreds of Watts (i.e., from 210 W at 5 GHz to 100 W for 2 GHz). From Fig. 5, for a 10 Gbps WDM system with 4 slices 3898.4 pJ per information bit is reported. This results in 3.89 W power consumption showing the potential advantage for power savings of the proposed approach. However, this power savings come at higher latency penalties, 183.5 μs.

IV. Conclusion

Scalable sliceable transceivers represent a novel solution to cope with the current demands increase. This paper presents, for the first time, a detailed description on the additional digital signal processing required at the transmitter and at the receiver side to slice and reconstruct the signal, respectively, as well as the power consumption required for each process. The analysis has been performed for both WDM and TDM systems. For 1 Gbps signals WDM outperforms TDM systems offering between 19.43% and 69.54% power consumption reduction for different sampling frequencies, where the latter is associated with the Nyquist sampling frequency. Using WDM also results in 65% DSP latency reduction. Additional latency performance penalty needs to be considered in TDM system accounting for the time that takes for each slice to reach the destination. For 10 Gbps signals WDM and TDM systems offer similar performance.

Considering the complete system, results show that power savings can be obtained by mitigating the bandwidth requirements of electrical and optical equipment when using sliceable transceivers. DSP power consumption of 1 Gbps NRZ signal with 4 slices at a 2 GHz sampling frequency is below 280 mW, whereas power savings in terms of electrical and optical equipment (e.g. the arbitrary waveform generator, AWG) is few tens of Watts (30 W when reducing from 120 MHz to 30 MHz bandwidth requirements). This power savings are obtained while achieving low latency penalties, i.e., 105 ns. For 10 Gbps DSP power consumption contribution for 4 sliced signal is 3-89 Watts leading to power savings up to hundreds of Watts from the electrical and optical equipment (e.g. 110 Watts from 5 GHz to 2 GHz AWG bandwidth requirements). For higher bit rates, power savings come at the expense of higher latency, 183.5 μs.

The results presented in this paper indicate that the benefits obtained by re-using low bandwidth equipment are not only relevant from cost point of view (CAPEX), but also from power consumption point of view, thereby influencing the operational expenditures (OPEX). Further power consumption reduction can be achieved by minimizing the energy to process a bit of information, for example exploring higher level modulation formats (e.g. duobinary, M-PAM, etc). Signal slicing creates a new migration strategy to supply high capacity channels (e.g. 10, 40, and 100 Gbps) in optical access networks at a cost-effective and energy-efficient manner.

Acknowledgment

The research leading to these results was partly supported and funded by the MARIE CURIE FENDOI and the MARIE CURIE ABACUS.

References

[1] ITU, *The world in 2014: ICT facts and figures*, ICT statistics, 2014.

[2] Cisco inc, *Cisco Visual Networking Index: Forecast and methodology, 2013-2018*, June 2014.

[3] E. Palkopoulou, D. A. Schuple, and T. Bauschert, *Quantifying CAPEX savings of homing architectures enabled by future optical network equipment,* Telecommunication systems, vol. 52, n.2, pp. 1123-1129, February 2013.

[4] S. Spolitis, C. Wagner, J. J. Vegas Olmos, V. Bobrovs, G. Ivanovs, and I. Tafur Monroy, *Experimental demonstration of a scalable sliceable transceiver for optical access network*, ACP/IPOC, 2014.

[5] D. J. Geisler, N. Fonatine, R. P. Scott, T. He, L. Paraschis, O. Gerstel, J. P. Heritage, and S. J. B. Yoo, *Bandwidth, scalable, coherent transmitter based on parallel synthesis of multiple spectral slices,* Proc. OFC/NFOEC, OTuE3, 2011.

[6] C. Lee, Y. Kim, and J. K. Rhee, *Green IP over WDM networking design considering energy-load proportionality*, ICNC, 2012.

[7] S. Saldaña Cercós, L. C. Resendo, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, *Power-aware multi-layer translucent network design: an integrated OPEX/CAPEX analysis*, Proc. OFC, 2014.

[8] Tektronix, *Techniques for extending real-time oscilloscope bandwidth,* White paper tektronix 2014.

[9] S. J. Savory, *Digital signal processing for coherent systems*, Proc. OFC, 2012.

[10] P. Duhamel, *Implementation of "Split-Radix" FFT algorithms for complex, real, and real-symetric data,* Transactions on acoustics, speech, and signal processing, vol. 34, n. 2, pp. 285-295, April 1986.

[11] H. V. Sorensen, M. T. Heideman, and C. Sidney Burrus, *On computing the split-radix FFT,* Transactions on acoustics, speech, and signal processing, vol. 34, n. 1, pp. 152-162, February 1986.

[12] B. S. G. Pillai, B. Sedighi, K. Guan, N. P. Anthapadmanabhan, W. Shieh, K. J. Hinton, and R. S. Tucker, *End-to-end energy modeling analysis of long-haul coherent transmission systems*, Journal of Lightwave Technology, vol. 32, n. 18, September 2014.

[13] online available: *http://literature.cdn.keysight.com/litweb/pdf/5990-7516EN.pdf*. September 2014.

[14] online available *http://literature.cdn.keysight.com/litweb/pdf/5991-0364EN.pdf*. September 2014.

[15] online available *http://literature.cdn.keysight.com/litweb/pdf/5991-3272EN.pdf*. September 2014.

[16] online available *http://literature.cdn.keysight.com/litweb/pdf/5991-0692EN.pdf*. September 2014.

Paper 2: Power-Aware Multi-Layer Translucent Network Design: an Integrated OPEX/CAPEX Analysis

Silvia Saldaã Cercós, Leandro C. Resendo, Moisés R. N. Ribeiro, Anna Manolova Fagertun, and Idelfonso Tafur Monroy, "Power-aware multi-layer translucent network design: an integrated OPEX/CAPEX analysis," *OFC*, 2014.

Power-Aware Multi-Layer Translucent Network Design: an Integrated OPEX/CAPEX Analysis

Silvia Saldaña Cercós[1,*], Leandro C. Resendo[2], Moisés R. N. Ribeiro[3], Anna Manolova Fagertun[1] and Idelfonso Tafur Monroy[1]

[1] *DTU Fotonik, Technical University of Denmark, Kgs. Lyngby, Denmark * ssce@fotonik.dtu.dk*
[2] *Department of System Information at the Federal Institute of Technology of Espírito Santo, Vitória E. S. Brazil,*
[3] *LABTEL-Electrical Engineering Department, Federal University of Espírito Santo, Vitória E. S. Brazil*

Abstract: We propose a three-phase network design model minimizing CAPEX and OPEX in IP-over-WDM architectures. By forbidding reconfiguration (accounting for 58% of the OPEX) we achieve only 4.2% increase in power consumption at no CAPEX expenses.

OCIS codes: 060.4256, 060.4510

1. Introduction

The strong competition in the telecommunication market motivates network providers to satisfy traffic demands at the lowest cost possible. However, as traffic demands increase, with expected traffic volume to transcend 1.4 zettabytes by the end of 2017 [1], network designers are forced to evaluate the deployment of new technologies capable of higher data rates. One of the main drivers when designing a network architecture is cost. Network operators aim at both lower capital (CAPEX) and operational (OPEX) expenditures [2]. One possible solution to achieve less complex systems at lower costs is integration, thus, different techniques have been applied to support efficient consolidation of network layers [3]. Reducing deployed infrastructure and operational costs, by integrating network platforms has been presented in [4]. Along these lines, several works have focused on the analysis of IP-over-WDM networks, e.g. [5]. Additionally, current networks need to support heterogeneous and dynamic services, where traffic loads vary during the time of the day. Many research studies have focused on how to dynamically adapt the usage of the network resources to reduce power consumption (PC) and consequently, operational costs, e.g. [6]. The main contribution of this work is the establishment of an optimization model with three design phases for a two-layer architecture. The first design-phase defines the physical infrastructure (accommodating the peak traffic) by either minimizing the CAPEX cost or the energy cost, taking into account the daily price variations. The two last design-phases (for average and low traffic) deal with traffic engineering, via a re-optimization process inspired by [7], which seeks to reconfigure (re-route and re-groom) the traffic demands through the most power-efficient paths allowing part of the network equipment to be switched off. The network reconfiguration tasks impact on OPEX is also included.

2. Network model

We consider a translucent IP-over-WDM network architecture, where optical-electrical-optical (OEO) conversions at intermediate nodes are performed only when strictly necessary. This work is based on the modular Cisco CRS-3 router for the IP layer and the Cisco ONS 15454 MSTP for the WDM layer considering mixed line rates (MLR) transponders at 10, 40 and 100 Gbps [8]. Network infrastructure design is performed so as to accommodate the peak-hour traffic (first design-phase). We compare two different strategies for the first design-phase: 1) minimizing deployed equipment cost and 2) minimizing PC. Subsequently, in order to account for the dynamic traffic profile, where mean traffic and low traffic levels represent 80% and 40% of the peak-hour traffic respectively, traffic engineering is applied. This is achieved by using a subset network architecture from the first phase-design, avoiding, in this way, physical modifications of the deployed network, except by switching off unnecessary elements for properly meeting traffic variations. Therefore, traffic engineering is always aiming at reducing OPEX in our proposal. According to [1], network OPEX is composed of different components: certification and upgrades (6.7%), services cost (S-OPEX)(58.3%), network care cost (11.4%), space cost (10.2%), cooling cost (6.2%) and PC cost (P-OPEX)(7.2%). In this work, we present two of the contributors to OPEX and we analyze two possible strategies: 1) minimizing PC and 2) minimizing service cost (in terms of network reconfigurations). The latter is important because re-configurations require service disruption to reset the network to the optimal traffic engineering state [9]. The following mixed integer linear programming (MILP) formulation is used for the network dimensioning:

Given: *i and j*: Network nodes bearing connection *ij*; *s and d*: Source and destination node; *k*: Transponder capacity; *V[i][j]*: Adjacency matrix for virtual topology connection; *Traf[s][d]*: Traffic matrix; *IM and L*: Maximum number of interface modules and line cards per router chassis respectively; Cap^k: Traffic capacity of a transponder k; *S[k]*: Number of slots occupied by one transponder (1, 2 and 3 for 10, 40 and 100 Gbps respectively); $\delta_{ij,r}^{sd}$: Link-path indicator, this is a binary indicator that it is set to one in case the link *ij* is used in a route *r* to meet the demand *sd* and to zero otherwise. This indicator is provided by the Yen's algorithm; *V:* Maximum number of slots per transponder power chassis; $P_I, P_{RC}, P_L, P_{TC}, P_T^k$*:* Power consumption of an interface module, router chassis, router line card, transponder power chassis and transponder k, respectively; $C_I, C_{RC}, C_L, C_{TC}, C_T^k$*:* Normalized cost of an interface module, router chassis, router line card, transponder power chassis and transponder k, respectively.

Variables: $C_{r,w}^{sd}$*:* Amount of traffic in transceiver *k* using wavelength *w* and route *r* from source *s* to destination *d*.

X_{ij}^k *and* X_{ij}^{k*}*:* Transponders at *k* Gbps used in the link *ij* for the first, and for the other two design-phases, respectively.

T_{C_i}, R_{SRI_i}, R_{I_i}, R_{C_i} *and* R_{L_i}*:* Number of active transponder power chassis, short range interfaces (SRIs), interface modules, router chassis and line cards at node *i*.

SRI_{ij} *and* SRI_{ij}^**:* SRIs at 10 Gbps used in the link *ij* for the first and for the other two design-phases, respectively.

L_{ij} *and* L_{ij}^**:* Line cards at 140 Gbps used in the link *ij* for the first and for the other two design-phases, respectively.

T_i^k*:* Number of transponders *k* at node *i*.

Objective functions: Two objective functions have been defined: (1) for the first design-phase, minimizing CAPEX based on the normalized cost presented in Table 1. (2) for the second and third design-phases, minimizing the total network PC while routing all traffic demands.

$\sum_i (C_I \cdot R_{I_i} + C_{RC} \cdot R_{C_i} + C_L \cdot R_{L_i} + C_{TC} \cdot T_{C_i} + C_T^k \cdot T_i^k)$ (1) $\quad \sum_i (P_I \cdot R_{I_i} + P_{RC} \cdot R_{C_i} + P_L \cdot R_{L_i} + P_{TC} \cdot T_{C_i} + P_T^k \cdot T_i^k)$ (2)

Constraints:

$\sum_r \sum_k C_{r,k}^{sd} = \text{Traf}[s][d],\ \forall s,d$ (3) $\quad \sum_s \sum_d \sum_r \sum_k \delta_{ij,r}^{sd} \cdot C_{r,k}^{sd} \leq 10 \cdot SRI_{ij},\ \forall i,j;\ \text{V}[i][j]{=}1$ (4)

$\sum_s \sum_d \sum_r \delta_{ij,r}^{sd} \cdot C_{r,k}^{sd} \leq Cap^k \cdot X_{ij}^k,\ \forall k,i,j;\ \text{V}[i][j]{=}1$ (5) $\quad \sum_k Cap^k \cdot X_{ij}^k \geq 10 \cdot SRI_{ij},\ \forall i,j;\ \text{V}[i][j]{=}1$ (6)

$SRI_{ij} = SRI_{ji},\ \forall i,j;\ \text{V}[i][j]{=}1$ (7) $\quad X_{ij}^k = X_{ji}^k,\ \forall i,j,k;\ \text{V}[i][j]{=}1$ (8)

$R_{SRI} = \sum_j (SRI_{ij} + SRI_{ji}),\ \forall i,j$ (9) $\quad T_i^k = \sum_j X_{ij}^k,\ \forall i$ (10)

$R_{SRI_i}/14 \leq R_{L_i};\forall i$ (11) $\quad R_{SRI_i}/14 \leq R_{I_i};\forall i$ (12)

$(L+IM) \cdot R_{C_i} \geq R_{I_i} + R_{L_i},\ \forall i$ (13); $\quad V \cdot T_{C_i} \geq \sum_k \text{S}[k] \cdot T_i^k,\ \forall i$ (14); $\quad X_{ij}^k \leq X_{ij}^{k*}$ (15); $\quad SRI_{ij} \leq SRI_{ij}^*$ (16)

Eqn. (3) ensures that traffic demand requests are satisfied. Eqns. (4) and (5) account for the number of 10 Gbps SRIs and transponders, respectively, needed to satisfy each traffic demand. Eqn. (6) distributes the traffic from the SRIs on the set of available transponders to ensure OEO-conversions. Eqns. (7) and (8) are symmetry constraints (i.e. the number of SRIs and transponders from *i to j* is equal than from *j to i*). Eqns. (9-14) are used to limit the devices capacities in the network (i.e. physical limitations taking into account hosting constraints from the equipment). Eqns. (15) and (16) ensure that new devices are not employed for low and average traffic loads. Table 1 presents the PC and the normalized cost for the devices in the IP and the WDM layers [10]. From the link perspective, even though the optical in-line amplifiers are the dominant contributors we assume that they will always be on and have traffic load independent PC, similar to [5].

Table 1: IP and WDM layers power consumption and normalized cost [10]

IP-layer			WDM-layer		
Device	Cost	Power [W]	Device	Cost	Power [W]
Router chassis	61.5	7660	Power chassis	5	255
Fan	0.34	334	10 Gbps transponder	/	35
Route processor	0.85	215	40 Gbps transponder	1	130
Interface module	14	446	100 Gbps transponder	2.5	133
+ line card		+150	+ CFP module		+84

We evaluate our model on the 14-node nation-wide Deutsche Telekom network (DTN) with 23 bi-directional links. The study considers an expected increase in the total traffic volume of 23% per year [1], and we look at the optimal network architecture based on traffic demands for 2017.

3. Results and analysis

Fig. 1 presents the traffic profile in one of the nodes (solid black) [11] (similar traffic pattern is assumed for the entire network), and the electricity pricing (solid blue) [12] in Germany for 24 h. From Fig. 1 we estimate the P-OPEX contribution by multiplying the MWh consumed by every network configuration (low (green), average (grey)

and peak (red) traffic) with the price at that specific time. When using (1) in the first design-phase (i.e. minimizing equipment cost), a 3906.35 cost units CAPEX and 250.7 kW PC values are obtained. If no traffic engineering is applied this would result in 301.8 €/day P-OPEX. After performing traffic engineering, 209.8 kW for average and 139.4 kW for low traffic are obtained. Considering the daily price variations, this represents 243.85 €/day P-OPEX, which yields in 19% total cost savings compared to the always-on configuration. From service cost point of view, 340 reconfigurations are performed. When using (2) (i.e. minimizing power consumption) in the first design-phase a 4134.35 cost units CAPEX and 247.4 kW PC values are obtained, which represents a 6% increase in the overall CAPEX and 3% decrease in P-OPEX. The average and the low traffic PC are 209.2 kW and 140.0 kW respectively. From service cost perspective, 320 reconfigurations are performed, obtaining 6% reduction. A larger benefit can be observed if no reconfigurations are allowed when performing traffic engineering. For this scenario 222.7 kW and 148.4 kW PC values for the average and the low traffic are obtained, which implies only 4.2% increase on P-OPEX (253.4€/day), while maintaining the same CAPEX level (4134.35).

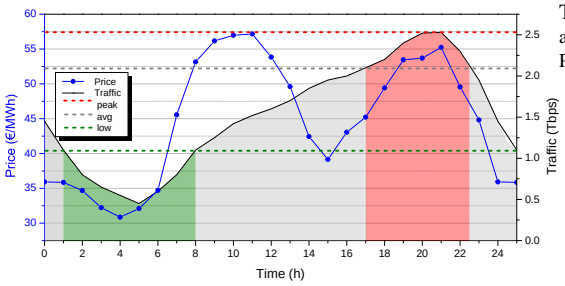


Fig. 1: Daily traffic variations (solid black) and price (circle blue).

Table 2: CAPEX and OPEX (for power consumption and service cost) values for three different strategies. PC: Power consumption, μ_r: Reconfigurations.

	Obj. function in phase 1		
	Eqn (1)	Eqn (2)	
	Traffic engineering in phase 2 and 3		
	$\mu_r \neq 0$		$\mu_r = 0$
CAPEX	3906.35	4134.35	4134.35
PC [€/day]	243. 8	242.9	253.4
μ_r	340	320	0

4. Conclusion

We presented an analysis for an IP-over-WDM network architecture considering OPEX, from power consumption and reconfiguration perspectives, and CAPEX. The paper presents for the first time an OPEX analysis broken down into two components: daily cost variations and re-configuration costs. A MILP model has been created to compare a network design aimed to minimize CAPEX cost versus a design aimed to minimize power consumption. Our results indicate that the benefits obtained by designing the network for minimum power consumption are negligible compared to designing for minimized CAPEX, since during the day electricity price does not follow the same profile as network traffic. The network design minimizing the power consumption yields only to a 6% increase in CAPEX, which suggests that more benefits can be achieved in the future by optimizing the network for power consumption since future trends indicate that new laws and regulations will introduce taxes proportional to the overall network power consumption. Moreover, we show that only 4.2% increase in power consumption OPEX is achieved by not allowing network reconfigurations, thereby improving the service OPEX, which accounts for more than 58% of the network OPEX.

References

1. Cisco, “Cisco visual networking index: Forecast and methodology, 2012-2017,” *Cisco white paper*, 2013.
2. R. Huellsermann et al. in *J. Opt. Netw.*, vol. 7, n. 9, Sept. 2008.
3. E. Palkopoulou et al, in *DRCN*, pp. 70-77, 2009.
4. F. Idzikowski et al. in *e-Energy'12*, 2012.
5. G. Rizzelli et al, in *Computer networks*, vol. 56, pp. 2443-2455, 2012.
6. A. Morea et al, in *J. Lightw. Technol.*, vol. 31, n. 5, March 2013.
7. L. Resendo et al, in *Photonics in Switching*, vol.1, no.3 pp.16-18, October 2006.
8. [Online] Available: *http://www.cisco.com*, 2013.
9. L. Resendo et al, *Journal of Optical Networks* vol. 7, pp. 533-549, 2008.
10. A. Klekamp et al, *J. Lightw. Technol.*, vol. 30, n.2, Jan. 2012.
11. [Online] Available: *http://www.de-cix.net/*
12. [Online] Available: *http://www.epexspot.com/*

Paper 3: A Power-Aware Rationale for Using Gross-Grained Transponders in IP-over-WDM Networks

Silvia Saldaña Cercós, Leandro C. Resendo, Moisés R. N. Ribeiro, Anna Manolova Fagertun, and Idelfonso Tafur Monroy, "A power-aware rationale for using gross-grained transponders in IP-over-WDM networks," in *JOCN*, 2015.

A Power-Aware Rationale for Using Coarse-Grained Transponders in IP-over-WDM Networks

Silvia Saldaña Cercós, Leandro C. Resendo, Moisés R. N. Ribeiro, Anna Manolova Fagertun, and Idelfonso Tafur Monroy

***Abstract*—Power consumption is becoming one of the most significant limitations while seeking for new solutions to cope with the traffic demand increase. 100 Gbps optical transmission technology has the potential to accommodate upcoming traffic demands with improved figures for W/Gbps compared to previous generations. However, the adoption of such coarse-grained bitrate granularity with lower flexibility for traffic grooming raises important questions: (1) which repercussions do they have on the overall power consumption and thus operational expenditures (OPEX) compared to legacy fine-grained designs (i.e., using 10 Gbps technology)? (2) What is the long term cost of coarse-grained designs? We define a power-aware Mixed Integer Linear Programming (MILP) formulation based on actual modular architectures where modules are upgraded as the network traffic increases. We introduce, for the first time, important practical constraints which were neglected by previous works, so that there is correspondence between interface modules and transponders, instead of assuming a single type of interface module per line card, and use accurate power consumption values instead of using approximations per port. We also present a comprehensive analysis on the trade-off between power consumption and available optical capacity, and power consumption and CAPEX for three different scenarios, defining the impact of provisioning the network with higher granularity transmission technology. Regarding the available capacity optical capacity vs. power consumption, 37.7% additional optical network capacity is achieved when using exclusively 100 Gbps technology at 13.3% to 32.4% power consumption expenses and 15.3% optical network capacity at only 2.6% to 8.8% power consumption penalty when using 40, and 100 Gbps technologies. From CAPEX perspective, up to 19.4% savings can be achieved by provisioning ahead using coarse-grained designs.**

***Index Terms*—Energy-efficiency, optical networks, traffic grooming.**

I. Introduction

As the Internet traffic continues to grow in volume and in diversity of required quality of service, the challenge for network architects is to design future-proof optical networks with enough capacity to cope with client traffic demands. A natural solution to maximize the use of fiber capacity is to deploy the latest transponder generation with higher aggregated bitrates, despite their unfavorable cost/benefit when compared with former transponder generations [1]. The challenge of designing optical networks with enough capacity to cater for traffic growth cannot be detached from energy efficiency considerations. Therefore, the subject of power consumption in core networks is an active research area. For example, a large number of analyzes document power consumption values for diverse optical networking equipment including Internet protocol/multiprotocol label switching (IP/MPLS), Ethernet, optical transport networking (OTN) and wavelength division multiplexing (WDM) technologies [2]–[4]. Other analyzes focus on energy-aware protocols [5], and trade-offs between quality of service (QoS) compliance and power consumption [6]. Furthermore, power-awareness has been evaluated both in IP and WDM layers and more recently multi-layer power-aware techniques were reported in [7]. Traffic engineering is an indispensable techno-economical tool for designing an efficient network that exploits the capabilities of the deployed resources. Along that line, emerging work is focusing on IP-over-WDM to identify the most energy-efficient architecture among opaque (where optical-electronic-optical (OEO) conversion is performed at each intermediate node), transparent (where optical signals are not OEO-converted at each intermediate node), and translucent (where OEO-conversion is performed only at those intermediate nodes where re-grooming is needed) as reported in [8].

Optimization models can provide network designers with the best combination of network elements for a given economic (e.g. reduce capital expenditures (CAPEX)), and/or technical (e.g. reduce power consumption or increase network capacity) goal. In [9], and [10] mixed integer linear programming (MILP) models are reported focusing on traffic engineering to dynamically adapt the routing in the virtual layer and realization of lightpaths in the physical layer. These models aim at the minimization of the number of employed active components. In [11] a new formulation for energy-efficient traffic grooming of scheduled sub-wavelength demands is presented accounting for connection holding times to route the demands in a way that a maximum number of lightpaths can be switched off at a given time. Switching on and off network components points out the need for considering dynamic traffic demands behavior. Consequently, energy optimization in IP-over-WDM networks, considering night-time low traffic demands to achieve energy savings when employing schemes that follow daily traffic demand variations, have been reported in [12], [13], and [14].

Although several papers have focused on minimizing power consumption in different layers, the work found in literature either account for approximations on power consumption per

S. Saldaña Cercós, A. Manolova Fagertun and I. Tafur Monroy are with the Department of Photonics Engineering, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark, e-mail: (ssce@fotonik.dtu.dk).

L. C. Resendo is with the Department of System Information at the Federal Institute of Technology of Espírito Santo, Vitória E. S. Brazil, e-mail (leandro@ifes.edu.br)

M. R. N. Ribeiro is with LABTEL-Electrical Engineering Department, Federal University of Espírito Santo, Vitória E. S. Brazil

port [4], or consider a single interface module independent on the type of transponder used [15]. We introduce for the first time a model that accounts for accurate power consumption values and constructive constraints, so that the interface module matches the transponders capacity. The outcome of previous presented models is not reliable, leading towards misleading solutions since they are based on simplistic approximations regarding those constructive features of modular equipment. Additionally, the potential of using exclusively high capacity transponders (e.g., 100 Gbps transponders), defined as coarse-grained design throughout this paper, in terms of available capacity and deployment cost has not been analyzed. This paper extends previous work presenting a novel model that considers realistic and feasible configurations and it also presents a comprehensive analysis on the trade-off between power consumption and available optical capacity and power consumption and CAPEX for both coarse-grained and legacy fine-grained designs which are based on 10 Gbps technology.

The rest of this paper is organized as follows: Section II provides an overview of the multi-layer network architecture including different network strategies to satisfy traffic demands and the power consumption contributions of each network component at both the IP and the WDM layers. Section III presents the power efficiency design and the mathematical model. Numerical results and network performance are described in Section IV. Section V highlights concluding remarks and future trends.

II. Network architecture: IP-over-WDM

We consider an IP-over-WDM network architecture justified by the dominance of IP traffic in optical backbone networks. Fig. 1 presents the two-layer network model under consideration where optical switching is performed by optical cross

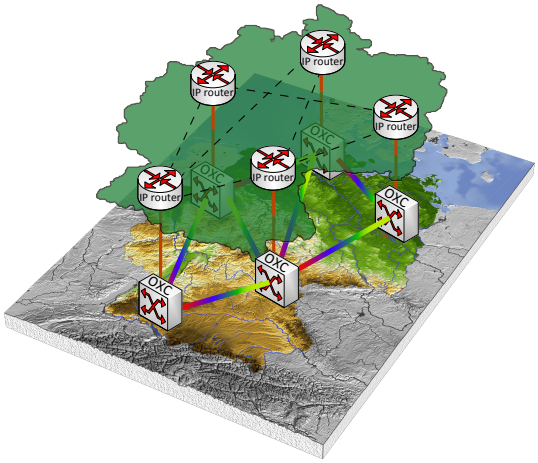


Fig. 1: IP-over-WDM model. Optical switching at OXC and IP traffic processing at the IP router. OXC, optical cross connect; IP, internet Protocol.

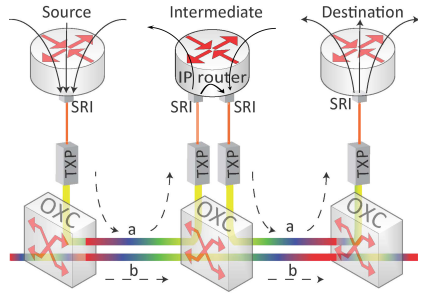


Fig. 2: Network architectures: path a) opaque, b) transparent. OXC, optical cross connect; SRI, short reach interface; TXP, transponder

connects (OXCs), physical links are realized by optical fibers and traffic processing is performed at the electronic level in the IP routers. In order to compensate for fiber attenuation and the insertion loss of OXCs, in-line optical amplifiers are considered (not represented in Fig. 1 and Fig. 2 for clarity). In the upper layer, the electrical domain, the dash black lines represent the logical topology between the IP routers. The electrical and optical domains are interconnected by a set of transponders (TXPs) which perform the necessary optical-electrical (OE) and electrical-optical (EO) conversions. Fig. 2 presents the two-layer connectivity at a node level. Communication between source and destination node can be implemented in various ways. Three different architectures are possible: transparent, opaque, and translucent. The power consumption of optical networks has a strong dependence on the architecture employed, therefore further details on these network architectures are presented in Section II-A.

A. Multi-layer networking strategies

To satisfy traffic demands between source and destination nodes, the IP router in the source node aggregates all the traffic demands onto the short range interface (SRI) capacity (black solid lines in Fig. 2). Afterwards, EO-conversion is performed for data transport through the optical channel to the destination node. In the destination node OE-conversion translates the data back into the electrical domain for further processing by the IP router. This communication can be accomplished in three different ways. In transparent networks (path *b* in Fig. 2), OEO-conversion in intermediate nodes is not present, thereby reducing electronic signal processing. However, physical layer impairments may limit transmission reach requiring signal regeneration at some intermediate node before reaching the destination node [15]. Opaque implementation, conversely, performs OEO-conversion at every transit node (path *a* in Fig. 2), allowing grooming, wavelength conversion and regeneration at the expense of extra electronic processing, power consumption and latency. Translucent networks are the third alternative approach where OEO-conversion and signal processing are performed only when strictly necessary. Translucent networks are a cost-effective and energy-efficient approach, where multi-layer design accounts for benefits of going to the IP layer for

traffic aggregation, achieving higher bandwidth utilization and reduced network costs [16]. Moreover, a translucent approach facilitates the use of switch on/switch off techniques in IP-over-WDM networks by allowing for re-packing demands into higher capacity lightpaths, effectively increasing the energy-efficiency of the used devices (i.e., the Watts per Gigabit per second [W/Gbps]). Consequently, the case study presented in this work considers a translucent network when analyzing the network power consumption.

B. Power consumption in commodity modular equipment

The analysis presented in this article is based on a modular configuration of the IP and the WDM layers, where scalability is achieved by interconnecting extra modules in the node. This modular design is based on provisioning the network with enough resources to allocate all traffic demands. Modular components with an analysis on the power consumption of each element are defined in this Section to present a comprehensive and realistic analysis on this important component of the network OPEX.

In the IP layer, data-plane processing tasks to handle network traffic flows such as traffic extraction (at the destination node), traffic aggregation (at the source and intermediate nodes), and grouping of low traffic demands onto larger capacity lines are considered. Fig. 3a) illustrates the modular configuration based on single chassis configuration for a Cisco CRS-3 router [18]–[20]. Cisco technology has been adopted for this analysis since power figures are available and it is widely deployed worldwide. Each router is composed of one router chassis interconnecting a maximum of 16 slots. If more than one chassis per node is required, these can be interconnected by an interconnect fabric (IF). In this chassis line cards at 140 Gbps throughput for the client side are inserted. Interface modules (IMs) are required since the communication between the IP and the WDM layers is done

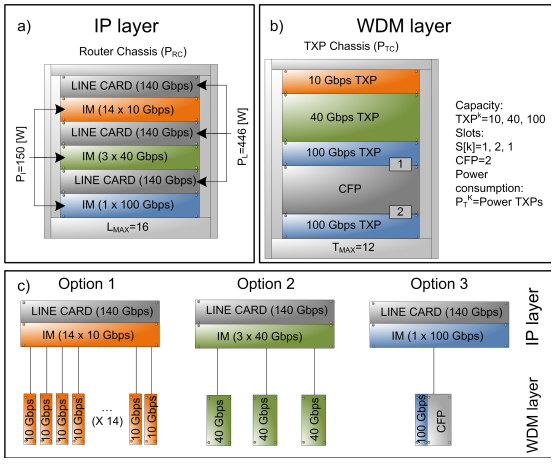


Fig. 3: IP and WDM modules. CFP, C form-factor pluggable; IM, interface module; IP, internet protocol; TXP, transponder; WDM, wavelength division multiplexing

TABLE I: IP and WDM layers power consumption. IM, interface module; IP, internet protocol; txp, transponder; WDM, wavelength division multiplexing

Device	*Power[W]*	Slots	*Device*	*Power[W]*	Slots
IP-layer			WDM-layer		
Router chassis	5700 [17]	16	Power chassis	55 [21]	12
Fan	334 [14]	0	10 Gbps txp	50 [22]	-1
Route processor	215 [14]	0	40 Gbps txp	130 [23]	-2
Line card	446 [18]	-1	100 Gbps txp	133 [24]	-1
IM	150 [19], [20]	-1	CFP line card	84 [25]	-2

through a physical interface module. The interface modules have a correspondence 1:1 with the optical layer, which is represented in Fig. 3b). They are interconnected via the SRI elements illustrated in Fig. 2. Our mathematical model differs from other models presented in literature in that it is the first one accounting for realistic practical arrangements instead of approximations on power consumption per port [4]. Different interface module per transponder capacity are based on state-of-the-art commercial available equipment [19] and [20]. Thus, to interconnect the IP line cards with the 10, 40 and 100 Gbps transponders, interface modules of 14 SRIs at 10 Gbps (14x10), 3 SRIs at 40 Gbps (3x40), and 1 SRI at 100 Gbps (1x100) are used, respectively (see Fig. 3c)).

The WDM layer is equipped with OXCs. In this work, a typical OXC based on photonic matrix switching technology, which consists of optical amplifiers, multiplexers/demultiplexers and a number of transponders which perform the OEO-conversions, is considered. In the model presented in Section IV, pre-amp and booster amplifiers are not taken into account since the number of optical amplifiers per node may vary significantly among different node architectures. Moreover, for dimensioning the transponders in the WDM layer, the power consumption contribution of the optical amplifiers can be considered as static-base power consumption per node. As shown in Fig 3b) the modular approach of the WDM module consists of a power chassis which is provided with 12 slots for the transponders, and a number of 10 Gbps, 40 Gbps and 100 Gbps transponders which occupy 1, 2 and 3 slots from the power chassis, respectively. Note that the 100 Gbps module requires being combined with the C Form-Factor Pluggable (CFP) line card, which occupies 2 slots from the 3 slots previously mentioned, in order to provide 100 Gbps transponder capabilities [24], [25]. Power consumption values for each element and the number of slots that each device occupies are stated in Table I. A positive number of slots is used for the chassis, a negative number of slots is used for the devices that occupy some space in the chassis.

From the link perspective, optical in-line amplifiers are the dominant power consumption contributors. However, we assume that they have a constant power consumption independent on the traffic flow. Thus, for simplicity of the model optical in-line amplifiers are not accounted for.

III. Power efficient design

In this Section we detail the proposed design strategy and the mathematical modeling used to optimize the network power consumption. We present a design model for mixed optical layer configurations (i.e., mixed transponder rates) which determines the optimal transponder capacity to be deployed in terms of power consumption by defining routing tables and which network elements perform traffic grooming. The main goal of this approach is to provide a network configuration which ensures the optimal power consumption (i.e. the lowest) while satisfying peak-hour traffic demands. The present detailed model advances in relation to the work found in literature since they either account for approximations on power consumption per port [4], or considers a single interface module without taking into account viable practical configurations (i.e. correspondence 1:1 between the transponders and the interface modules employed) [15].

MILP models output the mathematically optimal configuration disregarding other alternatives, even though the penalty in terms of power consumption of using another configuration (e.g. higher capacity technology) might be marginal. This paper fills this gap by also constraining the problem to golden rules from network architects in designing future-proof optical networks: maximize the use of fiber capacity deploying the latest transponder generation with higher aggregated bitrates.

The mathematical model, as well as parameters taken into account to define the different configurations are presented in Section III-A.

A. Mathematical modeling

MILP is used to formulate and solve network design and traffic engineering problems optimizing for power consumption. This model, which accounts for translucent network with electronic grooming [26], [27], uses the following notation:

Notation:
i and j: Network nodes, bearing connection ij (physical or logical connection).
s and d: Source and destination nodes respectively.
k = {10, 40, 100}: Variable index representing transponders (TXPs) capacity, 10 Gbps, 40 Gbps and 100 Gbps.

Parameters:

V[i][j]:	Adjacency matrix for virtual topology connection
N:	Number of nodes in the network
Traf[s][d]:	Traffic matrix
L_{MAX}*:*	Maximum number of slots per router chassis
TXP^k*:*	Traffic capacity of a transponder k (10, 40 or 100 Gbps)
P^k*:*	Traffic capacity of a port k (10, 40 or 100 Gbps)
S[k]:	Number of slots occupied by one TXP at rate k, where 10 Gbps, 40 Gbps and 100 Gbps TXPs occupy 1, 2, and 1 respectively) (i.e. S[k]=(1,2,1))*
Y[k]:	Number of SRIs at rate k per interface module (IM), where 1 IM can support 14, 3, and 1 SRIs at 10 Gbps, 40 Gbps and 100 Gbps respectively) (i.e. Y[k]=(14,3,1))
Z:	Maximum number of slots per TXPs chassis
$\delta^{sd}_{ij,r}$*:*	Link-path indicator: binary indicator set to one when the link *ij* is used in a route *r* to meet the demand *sd* and to zero otherwise. It is provided by the Yen's algorithm [28]
P_I*:*	Power consumption of an IM
P_{RC}*:*	Power consumption of a router chassis
P_L*:*	Power consumption of a router line card
P_{TC}*:*	Power consumption of a TXPs chassis
P^k_T*:*	Power consumption of a TXP k
P_{CFP}*:*	Power consumption of a CFP

*Note that 100 Gbps TXP occupies 1 slot, but 2 extra slots are taken into account for the CFP module [24], [25].

Variables:

C^{sd}_r*:*	Amount of traffic using route *r* from source *s* to destination *d*
W^k_{ij}*:*	Number of ports at *k* Gbps in the link *ij*
X^k_{ij}*:*	Number of TXPs at *k* Gbps in the link *ij*
T^k_i*:*	Number of TXPs *k* at node *i*
T_{C_i}*:*	Number of TXPs chassis at node *i*
R_{I_i}*:*	Number of IMs at node *i*
R_{C_i}*:*	Number of router chassis at node *i*
$R^k_{P_i}$*:*	Number of ports at k Gbps at node *i*
R_{L_i}*:*	Number of router line cards at a node *i*
CFP_i*:*	Number of CFP at a node *i*

Objective Function:

The objective function seeks to minimize the total network power consumption while routing all traffic demands.

$$\sum_i (P_I \cdot R_{I_i} + P_{RC} \cdot R_{C_i} + P_L \cdot R_{L_i} + P_{TC} \cdot T_{C_i} + P_{CFP} \cdot CFP_i + \sum_k (P^k_T \cdot T^k_i)) \quad (1)$$

Constraints:
Equation 2 ensures that all traffic requests are fulfilled.

$$\sum_r C^{sd}_r = \text{Traf}[s][d],\ \forall\, s, d \quad (2)$$

Expression 3 accounts for the number of *k* Gbps ports from the IM needed for each traffic demand using the connection *ij*.

$$\sum_s \sum_d \sum_r \delta^{sd}_{ij,r} \cdot C^{sd}_r \leq \sum_k P^k \cdot W^k_{ij},\ \forall\, i, j;\ \text{V}[i][j]\text{=1} \quad (3)$$

Expression 4 accounts for the number of *k* TXPs needed to attend each traffic demand.

$$\sum_s \sum_d \sum_r \delta^{sd}_{ij,r} \cdot C^{sd}_r \leq \sum_k TXP^k \cdot X^k_{ij},\ \forall\, k, i, j;\ \text{V}[i][j]\text{=1} \quad (4)$$

Expression 5 distributes the traffic from the router line cards on the set of available TXPs through a set of available IM ports.

$$\sum_k TXP^k \cdot X_{ij}^k = \sum_k P^k \cdot W_{ij}^k, \ \forall\, i,j; \ \text{V[}i\text{][}j\text{]=1} \quad (5)$$

Equation 6 ensures the correspondence 1:1 illustrated in Fig. 3c) and it is directly derived from expressions 3, 4, and 5.

$$X_{ij}^k = W_{ij}^k \quad (6)$$

Equations 7 and 8 are symmetry constraints. They ensure that the number of ports in a router line card and TXPs used by the flows from node *i* to node *j* (W_{ij} and X_{ij}^k, respectively) are equal to the number of ports in a router line card and TXPs used by the flows arriving at node *j* from node *i* (W_{ji} and X_{ji}^k, respectively).

$$W_{ij}^k = W_{ji}^k, \ \forall\, i,j; \ \text{V[}i\text{][}j\text{]=1} \quad (7)$$

$$X_{ij}^k = X_{ji}^k, \ \forall\, i,j,k; \ \text{V[}i\text{][}j\text{]=1} \quad (8)$$

The following equations (9-15) are used to match the physical limits of the considered devices. Equation 9 accounts for the number of ports of type *k*.

$$R_{P_i}^k = \sum_j W_{ij}^k, \ \forall\, i,k \quad (9)$$

Expression 10 accounts for the number of IMs.

$$R_{I_i}^k \cdot Y[k] \geq R_{P_i}^k, \ \forall\, i,k \quad (10)$$

Equation 11 accounts for the number of router line cards. As illustrated in Fig. 3, the number of line cards is equal to the number of IMs.

$$R_{L_i} = \sum_k R_{I_i}^k, \ \forall\, i \quad (11)$$

Equation 12 ensures that the number of IMs is equal to the number of router line cars at each node, as illustrated in Fig. 3.

$$R_{I_i} = R_{L_i} \quad (12)$$

Equation 13 accounts for the number of TXPs at node i.

$$T_i^k = \sum_j X_{ij}^k, \ \forall\, i,k \quad (13)$$

Expression 14 ensures that each 100 Gbps TXP is connected to a CFP, where each CFP can serve maximum two 100 Gbps TXPs.

$$2 \cdot CFP_i \geq T_i^{100} \quad (14)$$

Expression 15 accounts for the number of router chassis.

$$R_{C_i} \geq \frac{R_{L_i} + \sum_k R_{I_i}^k}{L_{MAX}}, \ \forall\, i \quad (15)$$

And expression 16 accounts for the number of TXPs chassis

$$Z \cdot T_{C_i} \geq \sum_k S[k] \cdot T_i^k + 2 \cdot CFP_i, \ \forall\, i \quad (16)$$

IV. Network performance: case study

In this Section we report the obtained results. The results are divided in two parts. First, network design results on a small 3-node network with three different traffic scenarios are presented. This first set of results is presented in order to: 1) illustrate a small scale example to follow in details the modeling presented in this work, and 2) to verify and validate the model as the modular design, properly modeled, leads to results that contrast with the ones based on average power consumption. The analysis on the deployed nation-wide network that evolves according to future traffic demands predictions is presented and it is structured in three parts: 1) a network design analysis based on all the available technologies, 2) a coarse-grained design impact on power consumption by constraining the model to use higher capacity equipment, and 3) a detailed analysis on the available optical capacity for each configuration. To solve the MILP formulation CPLEX 12.2 was used in a computer with Corei7-2600 (3.4 GHz) and 8 GB Memory. Processing time was bounded in 2 hours, and 3.31% was the worst gap from the optimal outcome achieved. This small gap ensures the reliability of the results.

A. 3-node network illustrative example

The MILP model considers the equipment in both WDM layer and IP layer to find the most power-efficient design. Fig. 4 depicts an illustrative example of a 3-node bus topology with three traffic demands (dash blue lines). It also provides an overview of the four different network configurations

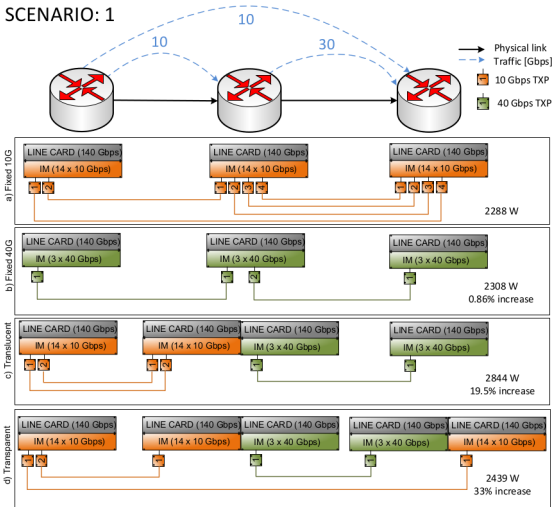


Fig. 4: Three-node network topology (solid black line) and traffic demands (dash blue lines) for a) single line rate (SLR) at 10 Gbps configuration, b) SLR at 40 Gbps configuration, c) mixed line rate (MLR) translucent configuration, and d) MLR transparent configuration

TABLE II: Physical design of 3-node network scenarios

Traffic scenario	Traffic matrix [Gbps]	Configuration node 1	node 2	node 3	TXP [Gbps]	Power consumption
Scenario 1	$\begin{bmatrix} 0 & 10 & 10 \\ 0 & 0 & 30 \\ 0 & 0 & 0 \end{bmatrix}$	2 - - -	4 - - -	4 - - -	10 40 100 CFP	2288 W
Scenario 2	$\begin{bmatrix} 0 & 40 & 10 \\ 0 & 0 & 30 \\ 0 & 0 & 0 \end{bmatrix}$	- 2 - -	- 1 - -	- 1 - -	10 40 100 CFP	2308 W
Scenario 3	$\begin{bmatrix} 0 & 40 & 10 \\ 0 & 0 & 110 \\ 0 & 0 & 0 \end{bmatrix}$	- - 1 1	- 3 1 1	- 3 - -	10 40 100 CFP	3598 W

considered by the model.[1] Fig 4 a) illustrates a single line rate (SLR) configuration, in this specific case at 10 Gbps, where transparent links are employed. The overall power consumption for this specific configuration is 2288 W. Fig. 4 b) depicts a SLR configuration, in this case at 40 Gbps, where the network aggregates traffic at the intermediate node. This second configuration provides overall power consumption is 2308 W which accounts for only 0.86% power consumption increase compared to the SLR configuration at 10 Gbps. Fig. 4 c) represents a translucent mixed line rate (MLR) configuration (i.e. with grooming) accounting for 2844 W power consumption, and Fig. 4 d) depicts a transparent mixed line rate (MLR) configuration consuming 2439 W, which represent 19.5% and 33% power consumption increase compared to the 10 Gbps SLR configuration, respectively.

A mistaken conclusion from current MILP models is lead by the fact that higher capacity technology is more power efficient (i.e. less W/Gbps, as presented in Table I) than lower capacity technology. Consequently, it is expected that high capacity technologies will outperform low capacity technologies, as translucent networks can do grooming by aggregating traffic into bigger pipes and resulting into lower power consumption. However, since the model presented in this paper considers actual practical constraints instead of approximations per port, it is possible to see in Fig. 4 c), that aggregating traffic could imply higher number of line cards and IMs leading to a higher overall power consumption. Additionally, results in Fig. 4 show that for the presented traffic demands, a 10 Gbps SLR configuration is the most power efficient. However, an SLR configuration with higher capacity (i.e. 40 Gbps) comes at only 0.86% power consumption expenses. These results indicate that over-provisioning the network can be done at a marginal OPEX increase in terms of power consumption.

Next, the model is evaluated under increasing traffic demands. Table II presents a sample of the resulting network configurations at a lowest power consumption under three different traffic demand scenarios. The traffic demand for each scenario is described by a traffic matrix, and the resulting power efficient configuration given by the MILP model is defined by the number of 10, 40, and 100 Gbps transponders at each node. Results show that with the increase of the demands the network migrates to MLR configuration, which is expected. Scenario 1 results in a low capacity SLR configuration, higher capacity SLR configuration is obtained for Scenario 2, and a translucent MLR configuration is obtained for Scenario 3, when the network design can benefit from traffic aggregation in the intermediate node. Although taking less W/Gbps, 100 Gpbs technology are only chosen when there is enough traffic to take most of its capacity (e.g., Scenario 3). Notice that each 100 Gbps transponder brings along a CFP module, which takes additional rack space. This may eventually lead to designs with more power-demanding chassis than the designs with fine-grained transponders (e.g., Scenarios 1 and 2).

[1] In Fig. 4 the chassis for both transponders and routers, as well as fan and route processor are not depicted for simplicity of the figure.

B. Nation-wide network scenario

In this Section the deployed nation-wide backbone network, Deutsche Telekom (DT) considered as a case study for the power consumption analyses is described. Fig. 5 illustrates the network topology which consists of 14 nodes and 23 bi-directional links between each neighbor nodes (one fiber per direction). The average link length is 186 km and the average node connectivity is 3.29.

In order to evaluate the network performance, the average traffic matrix presented in Table III is assumed [29].

Based on Cisco forecast [30], IP traffic is increasing 23% per year. Accounting for this 23% compound annual growth rate (CAGR) two different set of results are presented in this Section. First, an evolution of the most power efficient WDM equipment deployment, considering 10, 40, and 100 Gbps technologies, for average traffic demands from 2014 to 2020 is analyzed. The main goal of this analysis is define the desired equipment configuration based on future traffic demands, rather than seeing how a network provider upgrades the network based on the infrastructure already deployed. Therefore, the infrastructure configuration defined by the MILP model for a specific year is not constrained by the configuration that has been selected for previous years. The number of routes r for the link-path indicator was limited to 30 (i.e., the 30 shortest routes). The selected route by the model

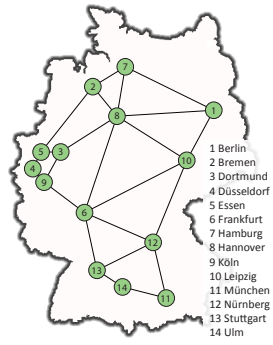


Fig. 5: Deutsche Telekom network topology

TABLE III: Deutsche Telekom average traffic matrix [Gbps] [29]

Node ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0.00	8.98	12.35	13.64	9.74	32.70	19.34	21.04	14.59	33.68	15.40	12.32	23.74	11.07
2	8.98	0.00	5.76	6.23	4.51	14.19	9.92	10.56	6.59	12.59	6.43	5.13	10.15	4.69
3	12.35	5.76	0.00	12.27	10.90	21.94	11.38	13.34	12.17	17.73	9.33	7.50	15.10	6.88
4	13.64	6.23	12.27	0.00	12.52	24.58	12.48	14.31	18.02	19.54	10.42	8.33	16.96	7.68
5	9.74	4.51	10.90	12.52	0.00	17.29	8.95	10.25	10.46	13.96	7.39	5.92	11.98	5.45
6	32.70	14.19	21.94	24.58	17.29	0.00	28.99	33.09	27.13	47.75	26.20	21.64	27.56	19.88
7	19.34	9.92	11.38	12.48	8.95	28.99	0.00	20.87	13.26	26.42	13.30	10.60	20.84	9.65
8	21.04	10.56	13.34	14.31	10.25	33.09	20.87	0.00	15.16	30.04	14.81	11.94	23.42	10.79
9	14.59	6.59	12.17	18.02	10.46	27.13	13.26	15.16	0.00	20.96	11.22	8.99	18.44	8.30
10	33.68	12.59	17.73	19.54	13.96	47.75	26.42	30.04	20.96	0.00	22.38	18.38	34.50	16.09
11	15.40	6.43	9.33	10.42	7.39	26.20	13.30	14.81	11.22	22.38	0.00	10.82	20.38	10.49
12	12.32	5.13	7.50	8.33	5.92	21.64	10.60	11.94	8.99	18.38	10.82	0.00	16.32	7.82
13	23.74	10.15	15.10	16.96	11.98	27.56	20.84	23.42	18.44	34.50	20.38	16.32	0.00	17.52
14	11.07	4.69	6.88	7.68	5.45	19.88	9.65	10.76	8.30	16.09	10.49	7.82	17.52	0.00

with the highest index number was 17. Longer index routes indicate longer routes physically, which in turn implies higher number of resources, thereby increased power consumption. It is expected that for a 14-node network, 30 routes are enough to achieve an optimal solution. Based on the initial results from a simple scenario from Fig. 4, where higher capacity SLR configuration comes at a low power consumption expense, a second set of results is presented in this Section, where only 40 and 100 Gbps technologies are considered. Results for an SLR configuration at 100 Gbps are also reported. Finally, an analysis on the optical network capacity when having a coarse-grained design (i.e. imposing the use of high capacity transponders) is reported.

1) 10, 40, and 100 Gbps technologies forecast: The first question we formulated was: which repercussions on the power consumption has the inclusion of new technologies such 100 Gbps technology?

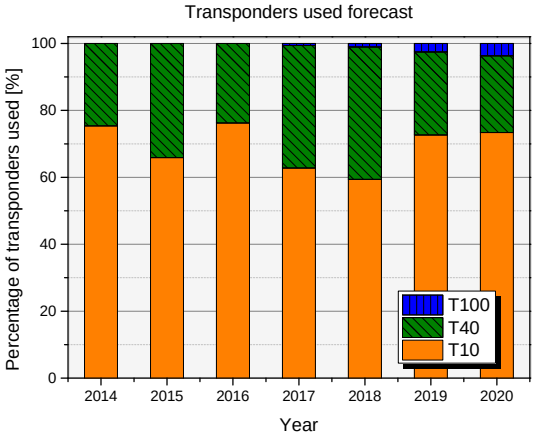


Fig. 6: Forecast for the use of transponders in the DT network with traffic matrix from 2014 to 2020. T10, 10 Gbps transponder; T40, 40 Gbps transponder; T100, 100 Gbps transponder

In order to answer this question, first we present an analysis on the considered network infrastructure for traffic demands from 2014 to a forecasted traffic demand in 2020 based on Cisco's predictions. The model selects the best configuration (i.e. the combination that offers the lowest overall power consumption) among all possible configurations using 10, 40, and/or 100 Gbps technologies. Fig. 6 shows that the dominant technology is 10 Gbps, and 100 Gbps technology only starts appearing in the overall picture in 2017 achieving below 10% deployment by 2020. 100 Gbps technology deployment remains below 10% due its constrained flexibility for traffic grooming. In contrast to already published results which show the potential for 40 Gbps technology for power savings [15] , with the model described in Section III, 10 Gbps technology is the dominant technology. This unexpected result is obtained due to the detailed physical considerations presented in the MILP formulation, showing that 10 Gbps technology utilizes better the network resources such as router and transponders chassis due to its granularity. The most power consuming element, as presented in Table I, is the router chassis, thus the outcomes of the model are the combination of resources that minimizes the number of line cards and interface modules needed to cope with the traffic demands. Minimizing the number of line cards is achieved by maximizing the use of interfaces per interface module. Based on the practical constraints presented in Fig. 3, in order to maximize the utilization of each interface module, the optimal outcomes are the ones able to allocate multiple of 14 and 3 number of 10 Gbps and 40 Gbps transponders per node, respectively.

Table IV presents the number of 10, 40, and 100 Gbps transponders used per node for three consecutive years (i.e. 2015, 2016, and 2017) as an numerical example. It shows how the node configuration for each year aims at deploying multiple of three 40 Gbps transponders, and multiple of 14 10 Gbps transponders. This results in the allocation of resources to achieve the most power efficient configuration

TABLE IV: Number of 10, 40, and 100 Gbps per node for years 2015, 2016, and 2017

			Node													
		TXP [Gbps]	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Year	2015	T10	14	10	14	14	14	14	11	14	14	12	11	14	14	0
		T40	6	3	3	5	2	12	6	9	6	12	6	3	9	6
	2016	T10	24	11	28	27	20	28	22	14	13	28	28	9	12	12
		T40	6	3	3	3	3	12	6	9	6	9	3	6	12	5
	2017	T10	12	14	28	14	8	28	13	12	14	24	12	13	13	11
		T40	12	3	3	9	6	12	12	12	9	12	9	6	15	6
		T100	0	0	0	0	0	1	0	0	0	1	0	0	0	0

for each year is shown in Fig. 6. The percentage of 40 Gbps technology fluctuates to accommodate traffic demands with a number of transponders multiple of three instead of following a (supposedly expected) linear increase.

2) 40 and 100 Gbps technologies forecast: In order to determine the impact on power consumption of deploying higher capacity technology (i.e., coarse-grained designs), a second set of results is presented where the model is constraint with using exclusively 40 and 100 Gbps technology.

Fig. 7 presents the percentage of 40 and 100 Gbps transponders used for the same network scenario. For this case study, 100 Gbps technology is present from current traffic demands, increasing its usage up to 37% in 2019.

As mentioned before, the outcome of the MILP models is the configuration that offers the mathematical optimized value. However, under single-variable objective function MILP solutions do not define how much better a solution is compared to an alternative solution where more constraints are applied. Fig. 8 presents the different power consumption values for the network based on the three presented configurations: Configuration 1 - MLR when all technologies are available (10, 40, and 100 Gbps-solid pattern), Configuration 2 - MLR when only 40, and 100 Gbps are available (diagonal stripes), and Configuration 3 - SLR when only 100 Gbps technology

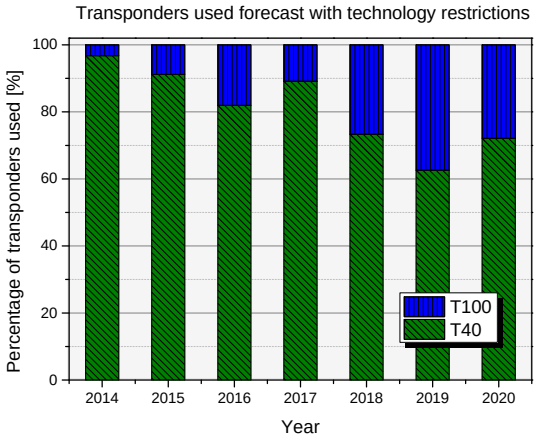


Fig. 7: Forecast for the use of transponders in the DT network with traffic matrix from 2014 to 2020. T40, 40 Gbps transponder; T100, 100 Gbps transponder

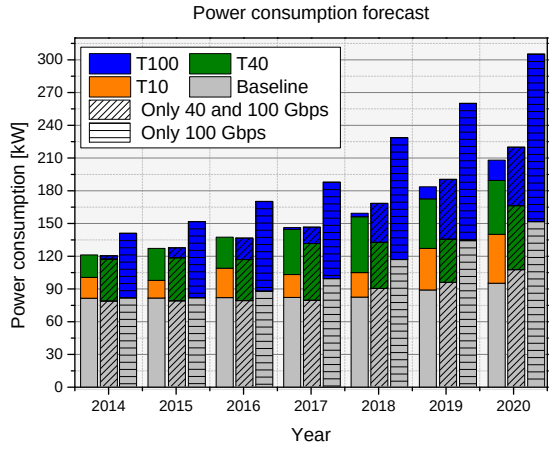


Fig. 8: Power consumption forecast forecast from 2014 to 2020. T10, 10 Gbps transponder; T40, 40 Gbps transponder; T100, 100 Gbps transponder; Other includes router and transponders chassis, fan, and router processor.

is available (horizontal lines). Power consumption results presented in Fig. 8 are also broken down into individual contributions for 10 Gbps technology (orange), 40 Gbps technology (green), and 100 Gbps technology (blue). 10, 40, and 100 Gbps technologies power consumption accounts for the power consumption of the transponders, interface modules, and line cards used for each technology. Router and transponders chassis, fan, and route processor power consumption contributions are accounted for in category "Baseline" (gray). The power consumption increase from Configuration 1 to Configuration 2 is below 5% until 2017 achieving only 2.6% increase in 2014, being 8.8% the maximum increase in 2020. In case of allowing the use of 100 Gbps technology exclusively, power consumption increases considerably resulting in 13.3% increase in the best case scenario in 2014, and up to 32.4% increase in 2020, and note that the heavier the traffic, the more power is proportionally consumed by the elements composing the baseline, in particular, due to extra chassis' space taken by the 100 Gbps technology.

As yearly traffic growth rates are assumed, 100 Gps lines are expected to be better utilized in later years, thus presenting lower relative power consumption. However, results show that the relative power consumption increases in 2020 as opposed to 2014. This outcome is explained by the fact that as the traffic increases, MLR configurations allow for better utilization of the line cards throughput (i.e., 140 Gbps). By using 14x10 Gbps transponders 140 Gbps throughput per line card can be achieved whereas when using 100 Gbps technology, only 100 Gbps throughput can be provided per line card. Consequently, using 100 Gbps technology, 40 Gbps per line card cannot be allocated for any traffic demand. Additionally, coarse-grained designs offer less flexibility in terms of grooming. As a result, additional line cards need to be installed with the corresponding interface module to meet

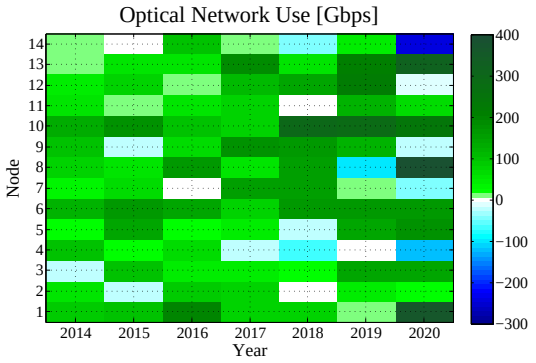


Fig. 9: Forecast of the optical capacity from Configration 2 in relation to Configuration 1

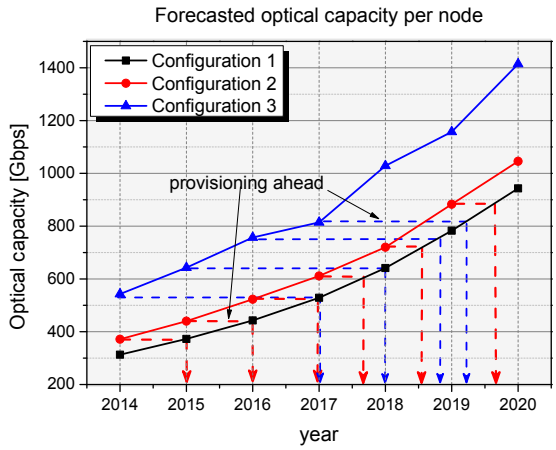


Fig. 10: Forecasted optical capacity per node for the three presented scenarios: for Configuration 1 (square, solid black), for Configuration 2 (circle, solid red), and for Configuration (triangle, solid blue)

traffic demands. Thus the increased number of router chassis (i.e., the most power demanding element) leads to a higher relative power consumption. The results from the reported MILP model show how minor increase in traffic brings along new devices and thus increased power consumption when implementing coarse-grained designs. This outcome was not possible to see with previous reported models where a single interface module per line card was assumed [4], [15], [31]. Higher power consumption penalty for 100 Gbps technology deployment is expected, since high capacity elements need to be used even for low traffic demands.

3) Total optical network capacity: The last question that was raised in the beginning of this paper was: what is the impact of a coarse-grained designed network? Considering that by using a mix of 40 and 100 Gbps technologies only 2-5% increase in the power consumption for the first five years, and below 8.8% until 2020 is achieved, a relevant question would be how much additional capacity per node can be obtained by deploying higher capacity components? In order to determine the obtained additional capacity, the term "optical capacity" is used to describe the maximum available capacity based on the deployed equipment in the network, e.g., if 10 Gbps technology is used to cope with a 10 Gbps traffic demand, the optical capacity is 10 Gbps. If 40 Gbps technology is used to cope with the same traffic demand of 10 Gbps, then the optical capacity is 40 Gbps since the deployed equipment allows for 30 Gbps additional traffic demands.

Fig. 9 presents the optical capacity comparison between Configuration 1 and Configuration 2. Results are presented in a three-dimensional space. The x-axis defines the year, the y-axis represents the nodes in the network, and the third dimension is given by the color. The year-node matrix represents the difference between the optical capacity per node per year between Configuration 2 and Configuration 1. Therefore, a green square in the node-year matrix indicates that the optical capacity for that node is higher when using only 40, and 100 Gbps technology, whereas a blue square stands for a higher optical capacity for that node when using a combination of 10, 40, and 100 Gbps technologies. White squares indicates equal optical capacity in both cases.

Results in Fig. 9 show that Configuration 2 outperforms Scenario 1 offering higher optical capacity up to 380 Gbps on node 8 in 2020. The average optical capacity increase is 81.42 Gbps per node which represents a 15.3% increase compared to the average optical capacity per node when 10, 40, and 100 Gbps technologies are available.

Fig. 10 illustrates the average optical capacity per node for the reference Configuration 1 (square, solid black), Configuration 2 (circle, solid red), and Configuration 3 (triangle, solid blue). The dashed blue arrows indicate at which point in time the reference configuration achieves the same average optical capacity per node as the 100 Gbps configuration. Results show that for the first two years the average optical capacity per node achieved by Scenario 3 is forecasted to be achieved three years after by the reference configuration. As the traffic increases, this time is reduced to two years. Analogously, the dotted red arrows illustrate at which point in time the reference configuration achieves the same average optical capacity per node as the Configuration 2. For the first three years, this capacity is achieved after a year, and as the traffic increases this time is reduced to 6-7 months.

4) CAPEX and power consumption trade-off: This Section presents the trade-off between CAPEX and OPEX in terms of power consumption. The CAPEX analysis described in this work is based on the IP and WDM layers equipment normalized cost values summarized in Table V which were reported in [14]. The methodology used to evaluate CAPEX is as follows: from Fig. 10 the optical capacity achieved by 100 Gbps technology (coarse-grained design) in 2014 (blue triangle) is first reached in 2017 by the reference configuration (black square) using 10, 40, and 100 Gbps technologies (legacy fine-grained design). The CAPEX associated to each of this curves is obtained by adding the cost of each of the devices

IP-layer		WDM-layer	
Device	Cost	Device	Cost
Router chassis	61.5	Power chassis	5
Interface module + line card	14	10 Gbps transponder	1
		40 Gbps transponder	2.5
		100 Gbps transponder + CFP module	3.75

TABLE V: IP and WDM layers normalized cost [31]

required for each configuration based on the MILP model output. The total CAPEX cost associated to Configuration 1 in 2017 (black square) is 2776 cost units whereas only 2236 cost units are obtained with the Configuration 3 in 2014 (blue triangle). This represents up to 19.4% CAPEX savings by provisioning ahead. These results are in accordance with [32] which states that the most cost-efficient solution is not always the least energy consuming one. It is important to note that: 1) 19.4% is an upper limit since cost variations as the technology becomes mature are not accounted for; and 2) these results are based on the average optical capacity per node. Using the average optical capacity per node does not ensure that the capacity is provided where this capacity is actually needed. Nevertheless, these results show future trends where coarse-grained designs will be more appealing for potential CAPEX savings and reduced power consumption as further module integration, thereby reduced footprint is implemented.

V. Conclusion

In this paper a realistic energy-aware MILP formulation for dimensioning an IP-over-WDM network has been provided. A model based on a modular configuration has been proposed where capacity upgrades are achieved by adding new modules accounting for the practical constraints both in the IP and the WDM layers.

The planning analysis has been extended by comparing power consumption of three different configurations: First, Configuration 1 when 10, 40, and 100 Gbps technologies are available, second, Configuration 2 when only 40 and 100 Gbps technologies are available, and third, Configuration 3 when solely 100 Gbps technology is available. Results show that 10 Gbps is the dominant technology when aiming at reducing power consumption since it facilitates maximizing the usage of each interface module, thereby reducing the number of required line cards. This result differs from previous reported models, since the model herein presented offers more detailed construction constraints showing how for coarse-grained designs minor increase in traffic requires additional power demanding equipment.To answer the question of which repercussions do coarse-grained designs have on the overall power consumption, the results obtained in this work show that Configurations 2 and 3 provide a power consumption increase compared to the Configuration 1 between 2.6% and 8.8%, and between 13.3% and 32.4%, respectively. On the other hand, using higher capacity technologies offers higher average optical capacity per node (i.e., available capacity for the deployed equipment). For the Scenario 2 an average of 15.3% increase in average optical capacity per node compared to the reference Configuration is achieved. 100 Gbps technology presents a 37.7% increase in the average optical capacity at a 13.3-32.4% higher power consumption penalty.

These results indicate the potential of coarse-grained designs for CAPEX savings, since network upgrades would be required in a longer period of time. Answering the question of what is the long term cost of coarse-grained design, the CAPEX and OPEX trade-off analysis presented in this paper indicates up to 19.4% CAPEX savings when using solely 100 Gbps technology, as network providers can provision 3 years ahead. This trade-off can be mitigated, making 100 Gbps solutions more appealing as integration for 100 Gbps interface modules emerges, reducing footprint and power consumption. Alternatively, a mix of 40, and 100 Gbps can be used to facilitate migration strategies towards higher capacity technologies at a low OPEX (i.e. power consumption) expenses.

MLR network architecture is shown to be cost effective in satisfying heterogeneous traffic demands [33]. This work focuses on traditional well-established WDM technology. The presented model can also be used to study the impact of improvements in individual modules on the network design. For future traffic growth flexible networking and flexible transceivers are an appealing alternative to MLR configurations to be considered.

Acknowledgment

The authors would like to thank Srikanth Thodupunoori technical marketing engineer at Cisco Systems for his guidance to understand the working logic of the modular system. The authors also acknowledge the Brazilian agencies CNPq, CAPES, and FAPES who partially supported this work.

References

[1] Rang-Chen Yu, *Cost efficient 100/400G Optical transceiver solutions*, Market Watch Panel, Oplink, OFC 2014.

[2] R. S. Tucker, *The role of optics and electronics in high capacity router*, in Journal of lightwave techonology, vol. 24, no. 2, December 2006.

[3] W. V. Heddeghem, F. Idzikowski, W. Vereecken, D. Colle, M. Pickavet and P. Demeester, *Power consumption modeling in optical multilayer networks*, Photonic Network Communications, vol.23, January 2012.

[4] W. V. Heddeghem, F. Idzikowski, E. L. Rouzic, J. Y. Mazeas, H. Poignant, S. Salaun, B. Lannoo and D. Colle. *Evaluation of power rating of core network equipment in practical deployments*, Online conference on green Communications (GreenCom), pp.126-132, 2012.

[5] B. G. Bathula and J. M. H. Elmirghani, *Green networks: energy efficient design for optical networks*, Wireless and Optical Communications Networks (WOCN), pp.1-5, 2009.

[6] J. Wang, S. Ruepp, A. Manolova, L. Dittmann, S. Ricciardi and D. Careglio, *Green-aware routing in GMPLS networks*, ICNC, 2012.

[7] G. Rizzelli, A. Morea, M. Tornatore and O. Rival, *Energy efficient traffic-aware design of on-off multi-layer translucent optical networks*, Computer Networks, vol. 56, no.10, pp. 2443-2455, 2012.

[8] F. Musumeci, D. Siracusa, G. Rizzelli, M. Tornatore, R. Fiandra and A. Pattavina, *On the energy consumption of IP-over-WDM architectures*, ICC, 2012.

[9] F. Idzikowski, S. Orlowski, C. Raack, H. Woesner and A. Wolisz, *Dynamic routing at different layers in IP-over-WDM networks-Maximizing energy savings*, Optical Switching and Networking, vol.8, no.3, pp.181-200, July 2011.

[10] F. Idzikowski, S. Orlowski, C. Raack, H. Woesner and A. Wolisz, *Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios*, in Proc. ONDM, pp.1-6, 2010.

[11] Y. Chen and A. Jaekel, *Energy efficient grooming of scheduled sub-wavelength traffic demands*, in Proc. OFC, 2011, paper OWI1.

[12] Y. Zhang, M. Tornatore, P. Chowdhury and B. Mukherjee, *Energy optimization in IP-over-WDM networks*, Optical Switching Networks, vol. 8, no.3, pp.177-180, July 2011.

[13] A. Morea, O. Rival, N. Brochier and E. L. Rouzic, *Datarate adaptation for night-time energy savings in core networks*, Journal of Lightwave Technology, vol. 31, no. 5, pp.779-785, March 2013.
[14] S.Saldaña Cercós, L. C. Resendo, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, *Power-aware multi-layer translucent network design: an integrated OPEX/CAPEX analysis*, in Proc. OFC, 2014, paper W2A.53.
[15] G. Rizzelli, A. Morea, M. Tornatore and A. Pattavina, *Reach-related energy consumption in IP-over-WDM 100 G translucent networks*, in Journal of lightwave techonology, vol. 31, no. 11, June 2013.
[16] A. N. Patel, C. Gao, J. P. Jue, X. Wang, Q. Zhang, P. Palacharla and T. Naito, *Traffic grooming and regenerator placement in impairment-aware optical WDM networks,* in Proc. ONDM, pp.1-6, 2010.
[17] Cisco product data sheet: *cisco 12000 series 1-Port OC-48/STM-16 POS/SDH Line Card and Cisco 12016 Chassis*.
[18] Cisco product data sheet: *Cisco CRS-3 Modular Services Card (Line card)*, January 2014.
[19] Cisco product data sheet: *Cisco CRS-3 14-Port 10GE LAN/WAN-PHY Interface Module*, January 2014.
[20] Cisco product data sheet: *Cisco CRS-3 1-Port 100 Gigabit Ethernet Interface Module*, January 2014.
[21] Cisco product data sheet: *Cisco ONS 15454 Multiservice Transport Platform*, January 2014.
[22] Cisco product data sheet: *10-Gbps Full-Band Tunable Transponder Card for the Cisco ONS 15454 Multiservice Transport Platform*, January 2014.
[23] Cisco product data sheet: *Cisco ONS 15454 40 Gbps Enhanced FEC Full Band Tuneable Muxponder Cards*, January 2014.
[24] Cisco product data sheet: *Cisco ONS 15454 100Gbps Coherent DWDM Trunk Card*, January 2014.
[25] Cisco product data sheet: *Cisco ONS 15454 2-Port CFP Line Card*, January 2014.
[26] L. Resendo, M. Ribeiro, and J. Pires, *Optimal multilayer grooming-oriented design for inter-ring traffic protection in DNI multiring WDM networks,* Journal of Optical Networks vol. 7, pp. 533-549, 2008.
[27] L. Resendo, M. R. N. Ribeiro, *Complementary Optimal Approaches for Survivable WDM Mesh Network Design: Transceiver Deployment and Traffic Engineering,* Photonics in Switching, vol.1, no.3 pp.16-18, October 2006.
[28] Jin Y. Yen, *Finding the K shortest loopless paths in the network,* Management science, vol. 17, no. 11, pp. 712-716, June 1971.
[29] DICONET Project: *D.2.1: Definition of dynamic optical network architecture,* Version: 1.00, September 14, 2010.
[30] Cisco, *Cisco visual networking index: Forecast and methodology, 2012-2017,* Cisco white paper, 2013.
[31] A. Klekamp U. Gebhard, and F. Ilchmann, *Energy and cost efficiency of adaptive and mixed-line-rate IP over DWDM networks,* in Journal of lightwave techonology, vol. 30, n.2, pp. 215-221, 2012.
[32] E. Palkopulou, D. A. Schupke, and T. Bauschert, *Energy efficiency and CAPEX minimization for backbone network planning: Is there a tradeoff?,* in Advanced networks and telecommunication systems, pp. 1-3, 2009.
[33] I. Stiakogiannakis, E. Palkopoulou, and I. Tomkos, *On the cost efficiency of flexible optical networking compared to conventional SLR/MLR WDM networks,* ICTON 2013.

Silvia Saldaña Cercós received the BSc (2009) and MSc (2011) degree in Telecommunications Engineering from Universitat Politécnica de Catalunya (UPC) in Spain and the Technical University of Denmark (DTU) respectively. She has participated in the ICT-CHRON European research project. Since July 2012 she has been a PhD student in the Department of Photonics Engineering at the Technical University of Denmark focusing her research activity on energy efficiency in optical networks.

Leandro C. Resendo graduated in mathematic, he received the MSc degree in Informatincs and the PhD degree in Electric Engineering, from Federal University of Esprito Santo, Brazil. Since 2004 he focuses his research on optimization methods applied to optical network design. He is currently working as a professor in the Department of System Information at the Federal Institute of Technology of Esprito Santo.

Moisés R. N. Ribeiro was born in Vitória/ES, Brazil in 1969. In 1995, he joined the Department of Electrical Engineering of Universidade Federal do Espírito Santo, Brazil. He received the B.Sc. degree in Electrical Engineering from the Instituto Nacional de Telecomunicaões, the M.Sc. degree in Telecommunications from the Universidade Estadual de Campinas, Brazil, and the Ph.D. degree from the University of Essex, UK, in 1992, 1996 and 2002, respectively. Stanford University visiting Professor in the 2010-2011 academic year with the Photonics and Networking Research Laboratory. His research interests include fiber-optic communication and sensor devices, systems and networks.

Anna Manolova Fagertun received a PhD degree at the Department of Photonics Engineering from the Technical University of Denmark in 2010, a MSc degree in Telecommunication from the Technical University of Denmark in 2006, and a BSc degree in Telecommunication Technologies from the Technical University in Sofia, Bulgaria in 2004. She has been involved in several national and international research projects (EU FP7 NoE BONE, COST 293, COST 804 , HTF supported "The road to 100G Ethernet", GEANT 3). She is among the first to receive an Industrial Post-Doctoral grant from the Danish Advanced Technology Foundation for a joint project with TDC A/S, Denmark. She is currently a researcher at DTU Fotonik - Department of Photonics Engineering. Her main research interests are within the areas of network modeling and simulation, multi-domain/multi-layer network integration, GMPLS/ASON, network resiliency, energy efficiency, physical layer impairment aware control plane and next generation transport network technologies.

Idelfonso Tafur Monroy received the M.Sc. degree in multichannel telecommunications from the Bonch-Bruevitch Institute of Communications, St. Petersburg, Russia, in 1992, the Technology Licentiate degree in telecommunications theory from the Royal Institute of Technology, Stockholm, Sweden, in 1996, and the Ph.D. degree from Eindhoven University of Technology, Eindhoven, the Netherlands, in 1999. He is currently a Professor and the Head of the Metro-Access and Short Range Communications Group, Department of Photonics Engineering, Technical University of Denmark, Lyngby, Denmark. In 1996, he joined the Department of Electrical Engineering, Eindhoven University of Technology, where was an Assistant Professor until 2006. He has participated in several European research framework projects in photonic technologies and their applications to communication systems and networks. He is currently involved in the ICT European projects GigaWaM and EUROFOS and is the Technical Coordinator of the ICT-CHRON project. His research interests include hybrid optical-wireless communication systems, high-capacity optical fiber communications, digital signal processing for optical transceivers for baseband and radio-over-fiber links, application of nanophotonic technologies in the metropolitan and access segments of optical networks as well as in short-range optical-wireless communication links.

Paper 4: Empirical Multichannel Power Consumption Model for Erbium-Doped Fiber Amplifiers

Silvia Saldaña Cercós, Getulio E. Rodrigues de Paiva, Marco Colazza Argentato, Anna Manolova Fagertun, Juliano Rodrigues Fernandes de Oliveira, and Idelfonso Tafur Monroy, "Empirical multichannel power consumption model for erbium-doped fiber amplifiers, In submission.

Empirical Multichannel Power Consumption Model for Erbium-Doped Fiber Amplifiers

Silvia Saldaña Cercós[(1)], Getulio E. Rodrigues de Paiva[(2)], Anna Manolova Fagertun[(1)], Marcio Colazza Argentato[(2)], Juliano Rodrigues Fernandes de Oliveira[(2)], and Idelfonso Tafur Monroy[(1)]

[(1)] Department of Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark, (ssce@fotonik.dtu.dk)
[(2)] Research and Development Center in Telecommunications (CPqD), Rod. Campinas, km 118.5, Brazil

Abstract *First experimental power consumption analysis and model of single and multi-stage booster erbium-doped fiber amplifiers (EDFAs) accounting for channel dependency. Results show that 48% of the total power consumption is associated to the number of channels that are amplified simultaneously.*

Introduction

Current optical networks are based on a wavelength multichannel network in order to accommodate end-user traffic demands. The challenge of how to increase channel transmission rates in optical fibers has generated significant research interest, spurring new technologies that lead towards optical fiber transmission links with 400 Gb/s and 1 Tb/s speeds[1]. In addition, complex advanced modulation formats have been studied offering the spectral-efficiency required to achieve such high bit rates[2]. To make these new technologies compatible with long and medium transmission reach, amplification is used. Thus, for next generation high capacity wavelength division multiplexing (WDM) networks, erbium-doped fiber amplifiers (EDFAs) are fundamental components. EDFAs are one of the most energy demanding devices within the optical fiber transmission link, accounting for up to 14% of the overall energy consumption[3]. Therefore, it is of relevance to study and understand its power consumption. Studies on the overall energy-efficiency for backbone communication networks have been presented, providing: 1) analytical power models[3], 2) power consumption per costumer models for optical networks[4], and 3) power consumption values for individual network devices[3]. More specific works for EDFAs include generalized models[5], models from a thermodynamic point of view[6], and models for multi-stage EDFAs[7]. However, power consumption models account only for a static power consumption value for EDFAs[8]. Assuming no control on the EDFA power consumption based on the number of simultaneously amplified channels limits the achievable power savings for backbone networks. This study includes experimental validation of EDFA power consumption, showing its dependency on dynamic channel load, and introduces a mathematical model based on empirical data.

Experimental power consumption analysis

In this work two Padtec LightPad il600G standalone booster EDFAs are tested, consisting of a single and a double stage erbium-doped fiber (EDF), respectively. A continuous wave (CW) laser diode at 980 nm is used to pump each EDFA, and a gain-flattening filter (GFF) is implemented because, even under optimum pump conditions, the gain spectrum may not be uniform. Two optical isolators are used to avoid back reflections. A noise figure (NF) of 5.5 dB for -5 dBm input power has been measured for both devices. We experimentally evaluate the power contribution from each sub-component of the EDFA to the overall power consumption. The experimental set-up is illustrated in Fig. 1. 80 CW WDM 50 GHz spaced channels are used at the transmitter as input to the system. This is achieved combining 40 C-band 100 GHz spaced WDM channels, with 40 channels 100 GHz spaced channels at +50 GHz offset. A wavelength selective switch (WSS) is used to remove the excess noise from amplified spontaneous emission (ASE). The gain spectrum is flattened through an optical GFF incorporated in the EDFA[5]. Measured optical spec-

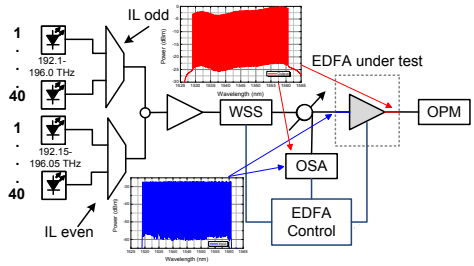


Fig. 1: Experimental set-up for a single- double-pump EDFA power consumption analysis

trums at the input (solid blue) and the output (solid red) of the EDFA are presented in the Fig. 1 as insets. The main goal of this experiment is to analyze the EDFA power consumption under channel loading variation. To achieve that, 1, 3, 12, and 50 channel loadings have been assessed. For the emulation of different channel loadings, -20 dBm power per channel is assumed, accounting for different input power into the EDFA for each channel loading. A variable optical attenuator (VOA) is used to adjust the power level injected into the EDFA under evaluation, emulating different channel loading conditions. The input power into the EDFA has been adjusted to -20 dBm, -15 dBm, -9 dBm, and -3 dBm input power for 1, 3, 12, and 50 channel loadings, respectively.

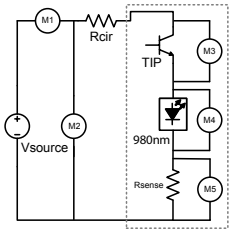


Fig. 2: Electrical control set-up for a single pump EDFA

As power consumption takes place in the electrical part of the EDFA control, five multi-meters (M_i, where $i = [1..5]$) are used to determine the different contributors of the EDFA electrical control set-up shown in Fig 2. The total current and voltage of the power supply is measured by M_1 and M_2, respectively. M_3 accounts for the collector-emitter bipolar junction transistor (TIP) power consumption. This transistor is used as a laser driver. M_4 accounts for the forward voltage of the pump laser, and M_5 accounts for the sense resistor power consumption contribution.

In order to experimentally evaluate the EDFA power consumption, four parameters are considered: 1) idle (i.e., static) power consumption, 2) channel loading variation impact on the overall power consumption (i.e., dynamic power consumption), 3) pump power required to achieve certain optical gain and its associated power consumption, and 4) power consumption based on the total optical output power. To define the static power consumption, the total circuitry contribution has been measured maintaining the pump laser disabled from the controller board. Results show an average power consumption of 6.34 W with a standard error of 0.02 W. Dynamic power consumption is evaluated by adjusting the input power into EDFA (i.e., varying the VOA) and its optical gain (i.e., increasing the pump laser power). Observations of the total power consumption as a function of the gain and channel loading, are presented in Fig. 3. The different gains are achieved by sweeping the pump laser power from 25 mW to 500 mW. At 25 dB gain the single-pumped EDFA consumed 7.4 W at single channel loading. Power consumption increases by 67% for 50 channels loading. This outcome is expected since lower channel loadings require lower pump rates to achieve higher gains. No power consumption difference is observed between single- and double-stage EDFA for a single channel loading. As the number of channels increases an increase on the power consumption for the double-stage EDFA is observed. At 20 dB gain, the marginal difference of 0.5 W between the powers consumed by a single- and double-stage EDFAs for the 50 channels loading case, founds a solid base for generalization of the empirical model for multiple-stage EDFAs. Results for different optical output powers with a fixed gain of 16 dB are presented in Fig. 4. 16 dB was chosen as a fixed gain since it is the necessary gain to compensate for fiber losses, anticipating 80 km of standard single mode fiber (SSMF) per span, and an attenuation of 0.2 dB/km. The optical output power curve shows a linear increase with a slope of only 1%. This indicates that no significant power consumption variance is experienced when using an EDFA with automatic gain control.

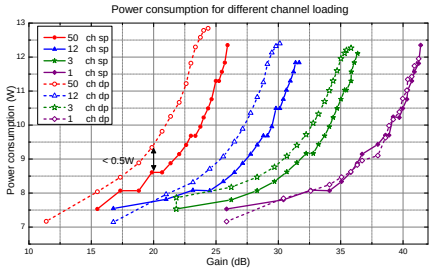


Fig. 3: Power consumption of EDFA for varied channel loading; ch: channel; sp: single-pump; dp: double-pump

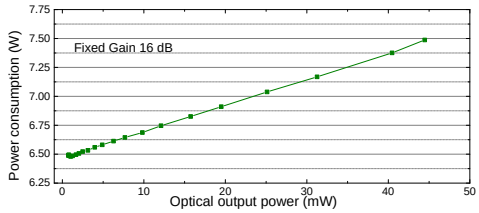


Fig. 4: Power consumption of the EDFA using fixed gain of 16 dB with the output power

EDFA power consumption model

We defined the EDFA control power consumption as in eq. (1), based on the measured data.

$$P_{EDFA} = P_C + K_p(n)(P_{pump} + P_r + P_{tip}) \quad (1)$$

Where P_{EDFA}, is the overall EDFA power consumption, P_C is the internal circuitry contribution when the EDFA is in idle state, K_p is a factor dependent on the pump power needed to obtain the desired gain which varies with the channel loading, n. Finally, P_{pump}, P_r, and P_{tip} are the contributions of the three electrical components of the EDFA control presented in Fig. 2.

A second formulation based on empirical data derived by well-known EDFA equations is also presented. Accounting for the NF of the EDFA, the signal-to-noise ratio (SNR) at the output of the EDFA can be expressed by eq. (2)[9].

$$SNR_{out} = \frac{P_{out}}{(2hv\Delta f + 4S_{sp}\Delta f)} \quad (2)$$

$$S_{sp} = (G-1)n_{sp}hv \quad (3)$$

Where hv is the photon energy, Δf is the detector bandwidth, and S_{sp} is the spectral density of spontaneous emission defined in eq. (3), being n_{sp} is the population inversion factor, and G the EDFA optical gain factor. The EDFA optical output power can be written as in eq. (4) based on the data from Fig. 3.

$$P_{out} = \frac{P_{EDFA} - 6.47}{0.01} \quad (4)$$

From eq. (2) and (4), we present in eq. (5) our second formulation for the EDFA power consumption, where α is described in eq. (6).

$$P_{EDFA} = \frac{SNR_{out}\alpha}{100} + 6.47 \quad (5)$$

$$\alpha = 2\Delta f(hv + 2S_{sp}) \quad (6)$$

To define the SNR required at the receiver side of a transmission link (SNR_{out}) to achieve certain bit-error-rate (BER), which modulation format is employed needs to be considered[10]. High order modulation formats have their constellation points closer spaced together intensifying the SNR requirements. From the SNR requirements curves for each modulation format presented in[10], the optical output power necessary to achieve a given BER and its associated power consumption can be derived from eq. (2) and (5), respectively.

Discussion and conclusion

As a conclusion, 52% of the EDFA overall power consumption comes from the control circuitry when the EDFA is in idle state. This circuit could be optimized to reduce the static power consumption of 6.34 W currently. The pump laser is the main component responsible for the variations of the overall power consumption. More channels loaded into the EDFA results in higher power consumption, since higher pump power is needed to achieve certain gain. 67% of the dynamic power consumption can be reduced by carefully adjusting the pump laser power. The presented analysis applies for booster single- and multi-stage EDFAs. The mathematical model proposed to analytically define the EDFA power consumption is based on measured data. It provides a good, generalized approximation for commercially available EDFAs. An EDFA power consumption model is essential step towards evaluating energy efficiency of next generation high speed optical networks. Thus, these results show high potential for global power savings while designing optical networks, by providing network designers with realistic power consumption values and model.

References

[1] H Takara et al., "120.7-Tb/s (7 SDM/180 WDM/95.8 Gb/s) MCF-ROPA un epeatered transmission of PDM-32QAM channels over 204 km," Proc. ECOC, Canes (2014).

[2] J.Estaran et al., "Experimental demonstration of capacity-achieving phase-shifted superposition modulation," Proc. ECOC, London (2013).

[3] W. V. Heddeghem et al, "Power consumption modeling in optical multilayer networks," in Photonic Network Communications, Vol. **24**, pp. 86-102, (2012).

[4] D. C. Kilper et al., "Power Trends in Communication Networks," in Selected Topics in Quantum Electronics, Vol. **17**, 2, pp. 275-284, (2011).

[5] Agrawal, Govind P., "Fiber-optic communication systems," Wiley (2010).

[6] S. Aleksic, "Energy and entropy flow in erbium-doped fiber amplifiers: a thermodynamic approach," in J. Lightwave Technol. Vol. **30**, pp. 2832-2838, (2012).

[7] A. W. Naji et al. , "Modeling and characterization of a new three-stage quadruple pass EDFA," ICCCE (2012).

[8] J. López et al., "Amplifier placements optimization for enhanced energy efficiency in optical transport networks," ICC (2014).

[9] G. Kweon, "Noise figure of optical amplifiers," in Journal of the Korean Physical Society, Vol. **41**, pp. 617-628, (2002).

[10] R.J. Essiambre et al., "Capacity limits of optical fiber networks," in J. Lightwave Technol. Vol. **28**, pp. 662-701, (2010).

Paper 5: Design of a Stateless Low-Latency Router Architecture for Green Software-Defined Networking

Silvia Saldaña Cercós, Ramon M. Ramos, Ana C. Ewald Eller, Magnos Martinello, Moisés R.N. Ribeiro, Anna Manolova Fagertun, and Idelfonso Tafur Monroy, "Design of a stateless low-latency router architecture for green software-defined networking," in *SPIE*, 2015.

Design of a stateless low-latency router architecture for green software-defined networking

Silvia Saldaña Cercós[a], Ramon M. Ramos.[b], Ana C. Ewald Eller[b], Magnos Martinello[b], Moisés R.N. Ribeiro[b], Anna Manolova Fagertun[a], and Idelfonso Tafur Monroy[a]

[a] DTU Fotonik, Technical University of Denmark, Kgs. Lyngby, Denmark;
[b] LabNERDS-Informatics and Electrical Engineering Department, Federal University of Espírito Santo (UFES), Vitoria E.S. Brazil

ABSTRACT

Expanding software defined networking (SDN) to transport networks requires new strategies to deal with the large number of flows that future core networks will have to face. New south-bound protocols within SDN have been proposed to benefit from having control plane detached from the data plane offering a cost- and energy-efficient forwarding engine. This paper presents an overview of a new approach named KeyFlow to simultaneously reduce latency, jitter, and power consumption in core network nodes. Results on an emulation platform indicate that round trip time (RTT) can be reduced above 50% compared to the reference protocol OpenFlow, specially when flow tables are densely populated. Jitter reduction has been demonstrated experimentally on a NetFPGA-based platform, and 57.3% power consumption reduction has been achieved.

Keywords: Energy-efficiency, optical networks, OpenFlow, NetFPGA -

1. INTRODUCTION

Current core network design strategies consolidate towards a cost-efficient solution by implementing a reduced number of core network nodes that support higher bit rates.[1] These strategies, combined with current Internet growth, lead to a key challenge for network operators, namely to cater to the large volumes of traffic. One solution for core networks is to define a forwarding plane of core nodes which ensures intra-domain routing packet transport while mitigating the complexity of network forward engines. Besides low-complexity for the elements in the forwarding plane, flexibility to cope with traffic changes on-demand is also required for meeting future requirements as they arise. Software defined networking (SDN) is a promising technology that offers this flexible and programmable networking platform with the control plane detached from the data plane.[2] Current SDN implementations focus on Ethernet switching primarily for data centers, that can benefit from centralizing the management intelligence.[3] However, some challenges arise when extending SDN for transport network architectures[4] , where scalability is on a major concern.[5] As the network scales up, the number of states related to the active flows in the network increases as well as the number of lookups table operations. This leads to an end-to-end higher latency.[6]

OpenFlow is the most used south-bound protocol within the SDN control architecture. However, it does not offer a simple design for forwarding engines, and it contributes to scalability problems due to the number of states related to the active flows. In this paper a protocol within the SDN architecture named KeyFlow for design of stateless low-latency router architecture is presented. KeyFlows aims at coping with latency challenges that arise in SDN transport network architectures since it does not require states associated to the active flows. KeyFlow reduces the complexity of the core forward engines by replacing the traditional table lookup by a simple operation: a division.[7,8] Consequently, latency linked to the table lookup bottleneck in the data plane of the SDN core networks is eliminated unlocking the potential usage of SDN technology in core networks. Besides latency, power consumption has been a concern in core networks design[9] since network equipment power consumption accounts for 14% of the overall information and communications technology (ICT) sector.[10] Ternary addressable content memories (TCAMs) have become an indispensable resource to provide the throughput and high packet

Further author information: (Silvia Saldaña Cercós E-mail: ssce@fotonik.dtu.dk)

processing rates that SDN requires. TCAMs are needed to perform the wildcard matching in flow table lookups despite that they are pricey and power-demanding.[11] Improvements upon the OpenFlow implementation in terms of TCAMs utilization have been presented in literature,[12] where per-port circuitry predicts the flow membership of a packet in order to bypass TCAM lookups. KeyFlow expands the state-of-the-art beyond reported research by investigating the potential benefits for a stateless low-latency and power-efficient router architecture.

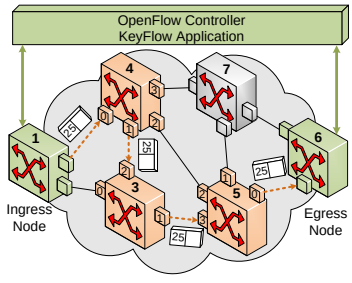
The rest of the paper is structured as follows. First, the operating principle for KeyFlow is detailed. Then, the design and implementation of a KeyFlow switch on a full line-rate 4-port 1 GbE NetFPGA-based is described. Third, latency and jitter results based on an open source platform, Mininet, for a KeyFlow prototype and OpenFlow switch are presented. Then an analysis on power consumption for a reference OpenFlow 1.0 switch is also presented. A comparison between OpenFlow 1.0 and a novel KeyFlow switch from power consumption point of view is performed. Finally, conclusions and future work is presented.

2. KEYFLOW: OPERATING PRINCIPLE

KeyFlow offers the possibility to reduce core network latency by removing the flow table lookup. Instead of the traditional flow table, KeyFlow is based on a simple division which concurrently reduces core network routers' complexity. The forwarding mechanism is based on the well-known residue number system (RNS).[13] The fundamental concept of RNS relies on the idea that every large number can be represented by a combination of small numbers obtained from the reminder of the division of this number by a set of co-primes. KeyFlow uses this mathematical property to define the path a packet follows. The large number represents the path ID, the set of co-primes represent the nodes' ID, and the reminder states the output port for each router.

Fig. 1 illustrates a KeyFlow network architecture example. A flow goes from the ingress node, $m_i = 1$, to the egress node, $m_i = 6$, through the intermediate nodes 4, 3, and 5, $m_i = (4, 3, 5)$. The path is illustrated with dash orange arrows. As depicted, the output ports are $p_i = (1, 1, 0)$ for each of the nodes, respectively. One of the benefits from KeyFlow is that the number of interactions between core nodes and controller is reduced. Communication occurs between node 1 and controller whenever there is no matching rule for a specific packet. The ingress node sends a request to the controller which computes a unique path ID based on the route that the packets needs to follow to reach its destination (e.g., the egress node with ID 6). In the illustrative example from Fig. 1a) this path ID is 25. This path ID is the so-called key. The controller sends a rule to the edge nodes with this information, so that every packet with the same destination is assigned with the same key. The ingress edge node assigns the key onto the packet header. KeyFlow core nodes do not have a flow table. Unlike traditional core intermediate nodes, KeyFlow core nodes perform a simple mathematical operation based on the key in the header of the packet and the node ID to determine the output port. The output port is determined by the reminder of the division between the key and the node ID.

Let $< X >_i$ be defined as the remainder from X/i. In the KeyFlow core node, X is the key at the packet's header and i is its own ID. In the example presented in Fig. 1a) node 4 performs the following operation to



(a)

Key calculation

m_i=(4, 3, 5): nodes ID
p_i=(1,1,0): output ports

1. $M = 4 \cdot 3 \cdot 5 = 60 \rightarrow M_1 = 15, M_2 = 20, M_3 = 12$
2. $L_1 = (M_1^{-1})_{\text{mod } m1} = (15^{-1})_{\text{mod } 4} = 3$
   $L_2 = (20^{-1})_{\text{mod } 3} = 2$
   $L_3 = (12^{-1})_{\text{mod } 5} = 2$
3. $X = (L_1 \cdot M_1 \cdot p_1 + L_2 \cdot M_2 \cdot p_2 + L_3 \cdot M_3 \cdot p_3)_{\text{mod } M}$
   $X = (45 + 45 + 0)_{\text{mod } 60} = 25$

(b)

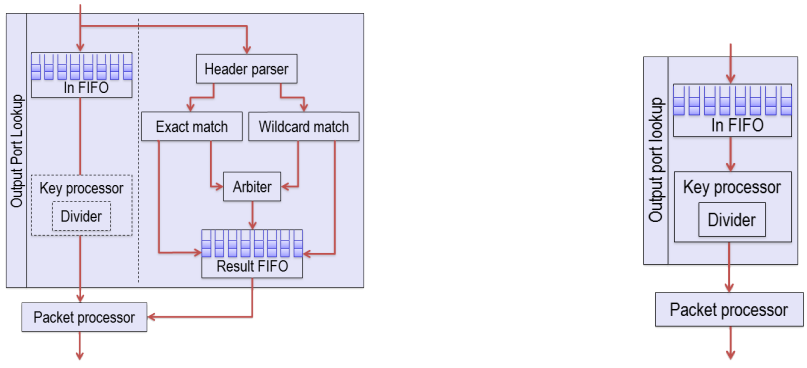
Figure 1: KeyFlow: Operating principle. 1a) KeyFlow network architecture example, and 1b) KeyFlow key calculation example

calculate the output port: $< 25 >_4 = 1$. Analogously, node 3 and 5 perform $< 25 >_3 = 1$, and $< 25 >_5 = 0$, respectively, to calculate their respective output ports. Fig. 1b) shows an example for the key calculation assessed in the control plane assuming that $m_i = (4, 3, 5)$ are the nodes IDs, 25 is the key, and $p_i = (1, 1, 0)$ are the respective output ports. Key calculation consists of three steps. 1) The parameter M_i is obtained by dividing M by each node ID, where M is the product of all nodes IDs. Thus, $M_i = M/m_i$. 2) The parameter L_i is calculated. L_i is the multiplicative inverse of M_i (i.e. $M_i \cdot L_i \equiv 1 \ (mod \ m_i)$). 3) The key, X, is calculated by summing up the product of each L_i, M_i, and output port for each node, $L_i \cdot M_i \cdot p_i \ (mod \ M)$.[7]

3. KEYFLOW AND OPENFLOW IMPLEMENTATION

In this work, the OpenFlow and the KeyFlow design are implemented using NetFPGA as the implementation platform. The hardware design is based on the reference implementations from Stanford University.[14] The 4-port 1 BgE NetFPGA Xilinx Virtex-II Pro 50 has been selected based on the following requirements: a platform capable of supporting high data rates (near-application-specific integrated circuit (ASIC) performance);[11,15] flexible enough to allow the implementation of a completely new core network forward engine. Due to the NetFPGA platform capabilities, it has been used previously to implement an IP-less forwarding fabric named Bloom-filter based.[16]

The OpenFlow reference design implemented on the NetFPGA includes MAC/CPU queues, output port lookup, and user data path modules.[14] From the OpenFlow modular NetFPGA pipeline structure presented in the literature[14] the work presented in this paper is based exclusively on the optimization of forwarding plane. Consequently, only the module associated to output port lookup is detailed. However, latency and power consumption calculations take into account all the modules of the reference pipeline.[14] Fig. 2 presents the modular NetFPGA structure for 2a) a Hybrid switch which includes both KeyFlow and OpenFlow capabilities, and 2b) a KeyFlow switch. For a conventional OpenFlow core network node, when a packet reaches the node, the packet enters into the output port lookup in order to find which output port to forward it to. In the output port lookup module, the header parser extracts relevant information from the packet header. This relevant information is sent to the exact match and the wildcard match in order to find a flow that matches the required information and determines the output port. The exact match uses two hash functions to match an exact flow. This process is performed off-chip on the static random access memories (SRAMs). The wildcard match, on the other hand, uses the TCAMs built using the Xilinx SRL16e primitives.[14] The arbiter module determines which solution to use in case of getting a match from both modules: exact match has higher priority than wildcard match. The forwarding information or associated action selected by the arbiter is stored in the Result first in first out (FIFO) buffer and sent to the packet processor. Fig. 2b) illustrates how complexity is reduced when using a KeyFlow switch as an alternative to the reference OpenFlow switch. The entire lookup process is replaced



(a) (b)

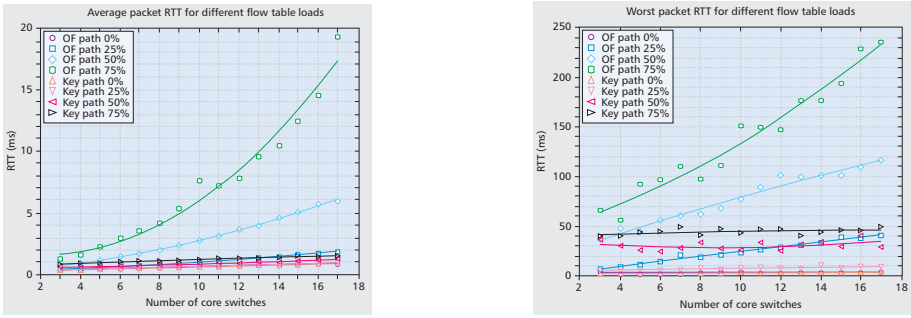
Figure 2: Modular NetFPGA pipeline structure for 2a) a Hybrid switch and 2b) a KeyFlow switch

by a single module which performs the division explained previously. By adding this module and replacing the traditional table look up three benefits are achieved: 1) latency is reduced. An OpenFlow switch by contrast offers a latency which is dependent on the flow table size, 2) jitter is also reduced. Under OpenFlow there is jutter because the delay associated with the table look up depends on where on the list the flow is placed, and 3) the need of using the pricey and power-demanding TCAMs is eliminated, since there is no need of performing wildcard match any longer, and instead a simple operation is done. In order to facilitate the migration towards KeyFlow core networks, a Hybrid switch has been implemented. A Hybrid switch keeps OpenFlow capabilities, but it has the possibility to enable the key processor module presented in Fig. 2a) so that the packet does not go through the header parser and subsequent modules. However, maintaining the modules gives the possibility to use the network fabric as OpenFlow for user convenience.

4. REDUCTION OF NETWORK LATENCY

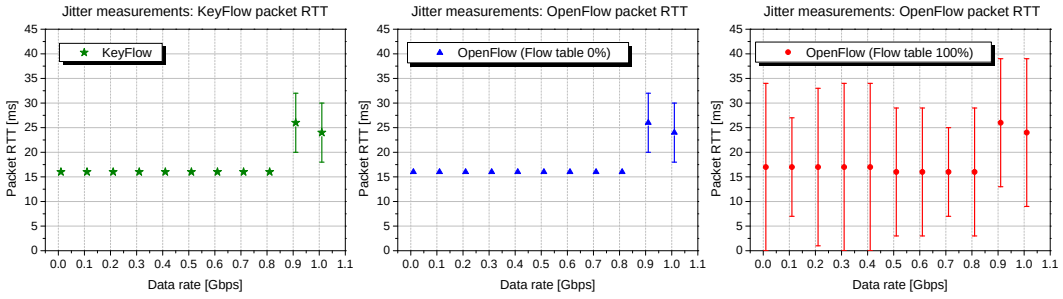
The scope of the presented latency analysis is to determine the switch latency associated with the table lookup bottleneck. This study gives an overview on the end-to-end latency impact when removing table look and replacing it with a simple mathematical operation. Tests have been done on a linear topology assuming that all the forwarding engines are under the same table utilization conditions, for table flow utilization of 0, 25, 50, and 75%. For the latency analysis the emulation platform Mininet has been used as an experimental testbed. For the experiment 5000 64-byte packets are generated and transmitted at constant time intervals. Fig. 3a) and Fig. 3b) show the average and the worse round-trip-time (RTT) for the packets for different flow table utilization as a function of the number of hops, in an emulated network with a reference OpenFlow and a KeyFlow switch. Results show that for a low flow table utilization (i.e., 25%) there is no significant difference on the packet RTT between the KeyFlow and OpenFlow switch. This is because the time needed to perform the division is equivalent to the time needed to execute the frequent cache hits for the lookup process. As the table utilization increases, the KeyFlow switch outperforms the OpenFlow switch. The OpenFlow switch offers an increased packet RTT with the number of hops, whereas the KeyFlow switch presents average values below 3 ms. There is a slight increase on the packet RTT when using a KeyFlow switch. This is expected, since higher number of flows requires additional lookup processing at the edge nodes. For the worse case scenario presented in Fig. 3b), values are one order of magnitude higher than for the average results presented in Fig. 3a). However, similar qualitative performance is observed in both scenarios.

In order to analyze the jitter impact for both an OpenFlow and a KeyFlow switch, the packet RTT has been measured on a NetFPGA-based platform for a single hop and 5000 packets for a) a KeyFlow switch (which does not have a flow table), b) an OpenFlow switch with the flow table utilization set at 0%, and c) the same OpenFlow switch with the flow table utilization set to 100%. As presented in Fig. 3 no differences in terms of



(a) (b)

Figure 3: RTT comparison between a reference OpenFlow and a KeyFlow switch: 3a) average and 3b) maximum packet RTT



(a) (b) (c)

Figure 4: Jitter measurements based on the packet RTT for 4a) a KeyFlow switch, 4b) an OpenFlow switch with the flow table utilization at 0%, and 4c) an OpenFlow switch with the flow table utilization at 100%

latency are expected for a single hop. Fig. 4a) and Fig. 4b) are in accordance with previous results, showing identical results for both the KeyFlow and the OpenFlow switch with an empty table flow. However, as the flow table utilization increases (Fig. 4), variations on the packet RTT are observed. This error bars represent the jitter impact on an OpenFlow switch since the packet RTT depends on where the rule associated to a specific packet is located on the table. Higher latency and jitter is achieved for bit rates close to the port line rate.

5. CORE FABRIC POWER CONSUMPTION

In this section power consumption measurements are presented. First, the used methodology is described. Then, a set of results is detailed including power consumption of a reference OpenFlow switch and a stateless KeyFlow switch. Power consumption is broken down into different components: 1) Total power consumption is presented accounting for both static and dynamic contributions, 2) Module-based power consumption is described including all the modules defined in the implementation section, and 3) A resource-based overview is given, where the power consumption of each of the NetFPGA resources is stated.

Measuring the power consumption of the two different switches is achieved by following three steps. The presented methodology is associated to the NetFPGA Xilinx platform. The first step includes a gate-level simulator, ModelSim. ModelSim is a Perl-based testing infrastructure which provides functional verification of the design. The outcome of the simulations includes the activity rates for each of the NetFPGA resources. These activity rates, also known as toggle rates, are needed to define how often there is a gate transition, and thus contributing into the overall power consumption. NetFPGA resources constitutes logic resources, signals, input/outputs (IOs), the input clocks for the block random access memories (BRAMs), and the digital clock managers. The simulation tool also provides the frequency for each of the internal clocks. Once these toggle rates are calculated they are exported into a .vsd file. Additionally, the design needs to be synthesized using a Xilinx ISE Virtex II 50 Pro NetFPGA. The synthesized design is described in a native circuit description (NCD) file. This file contains the logic of the design mapped to components. This file is also used to create a physical constraints file (PCF) which is important when determining the overall power consumption.[17] This synthesized design is used in the final step, together with the .vsd file to calculate the power consumption. The Xilinx XPower Analyzer (XPA) tool from Xilinx ISE design software is used. This tool provides the power consumption for each of the components based on the NCD and the toggle rates of each of the NetFPGA resources.

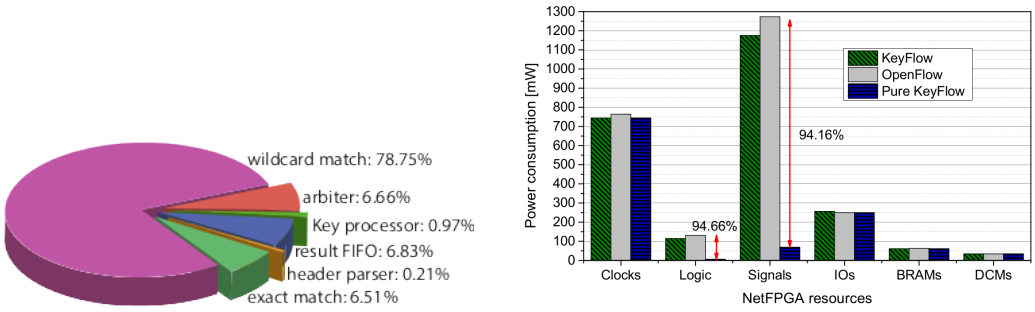
All the results presented in this section have been taken generating a single 1500-byte packet in order to avoid bottlenecks from buffering. Additionally, for the reference OpenFlow switch the flow table is set to 40 flows based on previous work presented in literature.[7]

Table 1 presents the static and the dynamic components of the power consumption for a) OpenFlow switch, b) Hybrid switch, and c) KeyFlow switch. The static power consumption is constant for the three implementations and it is measured to be 159 mW. From the dynamic power consumption perspective, 6% savings and 53.7%

Table 1: Total power consumption [mW]

	Static	Dynamic	Total
OpenFlow	159	2514	2673
Hybrid	159	2384	2543
KeyFlow	159	1164	1323

savings are achieved by the Hybrid and the standalone KeyFlow switch, respectively, compared to the reference OpenFlow switch. This outcome is achieved since the KeyFlow switch does not require the use of TCAMs when performing the table lookup. These results are verified when looking in detail at the contribution of each of the modules for the Hybrid switch. Fig. 5a) depicts a broken down analysis of each of the implementation modules. As expected, the most power-demanding module is the wildcard match since it requires TCAMs. It accounts for 78.75% of the overall power consumption of the output port lookup module. Exact match, result FIFO, and arbiter, which are used in the lookup process contribute with 6.51%, 6.83%, and 6.66%, respectively. The key processor accounts only for 0.97% of the overall power consumption. The data and control signals are the most power demanding resources for the NetFPGA as presented in Fig. 5b). Further analysis on the signals can be achieved by using a NetFPGA analyzer module which captures and allows designers to further work with *control*, *data*, in_{wire}, and out_{ready} signals.[18] A pure KeyFlow implementation, where the entire lookup process is eliminated as presented in Fig. 2b), achieves further savings compared to a KeyFlow switch that mantains the OpenFlow switch capabilities. This outcome is reflected in the resource-based power consumption analysis presented on Fig. 5b). Signals and logic power consumption can be reduced by 94.66% and 94.16%, respectively, compared to the reference OpenFlow switch. The reduction of the complexity of the forwarding engine leads to an overall power savings of 53.7%.



(a) (b)

Figure 5: 5a) module-based and 5b) resource-based power consumption contributions for an OpenFlow, KeyFlow (or Hybrid), and Pure KeyFlow switch

6. CONCLUSIONS

This article presents a new approach named KeyFlow for a south-bound protocol for SDN in order to offer a cost- and energy-efficient solution for the forwarding plane in transport networks. The proposal is based on an RNS. The presented results include RTT analysis evaluated on an emulation scenario based on a Mininet platform, a jitter comparison between an OpenFlow and a KeyFlow switch implemented on a NetFPGA-based platform, and power consumption comprehensive analysis for a reference OpenFlow switch, a KeyFlow switch, and a hybrid switch that offers both OpenFlow and KeyFlow capabilities. The RTT depends on the number of hops and how much populated the flow tables are. For flow table utilization below 25% no significant difference on the packet RTT is observed between the KeyFlow and the OpenFlow switch. For packets below 3 hops both switches perform similarly. As the number of hops increases KeyFlow outperforms OpenFlow specially for densely populated flow table switches, achieving more than 50% RTT reduction when using the novel KeyFlow

approach. KeyFlow is an alternative for current power-demanding high-performance switches. The major power consumption contribution in core network nodes are the TCAMs used for the wildcard match in the table lookup process. Wildcard match accounts for more than 78% of the output port lookup module, and thus, 57.3% power savings are achieved when replacing the traditional table lookup by a simple mathematical operation.

Future work includes embedded KeyFlow on wireless access points for mesh networks, and an enhanced version of KeyFlow to achieve higher power savings tackling the clock frequencies which are the second major power consumption contributor of a forwarding network node. On-chip solutions are the next step towards power-aware low latency SDN transport networks.

REFERENCES

[1] Betker, A., Gamrath, I., Kosiankowski, D., Lange, C., Lehmann, H., Pfeuffer, F., Simon, F., and Werner, A., "Comprehensive topology and traffic model of a nationwide telecommunication network," *Journal of Optical Communications and Networking* **6**, 1038–1047 (November 2014).

[2] Iovanna, P., Michele, F. D., Palacios, J. P. F., and López, V., "E2E traffic engineering routing for transport SDN," *Proc. OFC W1K.3* (March 2014).

[3] Marconett, D., Liu, L., and Yoo, S. J. B., "Optical FlowBroker: load-balancing in software-defined multi-domain optical networks," *Proc. OFC W2A.44* (March 2014).

[4] Gringeri, S., Bitar, N., and Xia, T. J., "Extending software defined network principles to include optical transport," *IEEE Communications Magazine* **51**, 32–40 (March 2013).

[5] Yeganeh, S. H., Tootoonchian, A., and Ganjali, Y., "On scalability of software-defined networking," *IEEE Communications Magazine* **51**, 136–141 (February 2013).

[6] Vencioneck, R. D., Vassoler, G., Martinello, M., Ribeiro, M. R., and Marcondes, C., "FlexForward: Enabling an SDN manageable forwarding engine in open vswitch," *10th CNSM and Workshop* (2014).

[7] Martinello, M., Ribeiro, M. R. N., de Oliveira, R. E. Z., and de Angelis Vitoi, R., "KeyFlow: A prototype for evolving SDN toward core network fabric," *IEEE Network* **28**, 12–19 (March 2014).

[8] Saldaña, S., Oliveira, R. E., Vitoi, R., Martinello, M., Ribeiro, M. R. N., Fagertun, A. M., and Monroy, I. T., "Tackling opeflow power hog in core networks with keyflow," *Electronic letters* (2014).

[9] Saldaña, S., Resendo, L., Ribeiro, M. R. N., Fagertun, A. M., and Monroy, I. T., "Power-aware multi-layer translucent network," *Proc. OFC W2A.53* (March 2014).

[10] Vishmanath, A., Sivaraman, V., Zhao, Z., Russell, C., and Thottan, M., "Adapting router buffers for energy efficiency," *CoNEXT (ACM)* , 19 (March 2011).

[11] Zerbini, C. A. and Finochietto, J. M., "Performance evaluation of packet classification on FPGA-based TCAM emulation architectures," *Globecom* (2012).

[12] Congdon, P. T., Mohapartra, P., Farrens, M., and Akella, V., "Simultaneously reducing latency and power consumption in OpenFlow switches," *IEEE ACM transactions on networks* **22** (June 2014).

[13] Garner, H. L., "The residue number system," *Trans. Electronic Computers* , 140–147 (1959).

[14] Naous, J., Erickson, D., Covington, G. A., Appenzeller, G., and McKeown, N., "Implementing an OpenFlow switch on the NetFPGA platform," *Proc ANCS* (2008).

[15] Jiang, W., "Scalable ternary content addressable memory implementation using FPGAs," *ANCS* (2013).

[16] Keinanen, J., Jokela, P., and Slavov, K., "Implementing zFilter based forwarding node on a NetFPGA," *Proc. of NetFPGA Developers Workshop* (2009).

[17] Xilinx, I., "FPGA editor guide," $http://www.xilinx.com/support/sw_manuals/2_1i/download/fpedit.pdf$ (November 2014).

[18] Goodney, A., Shailesh, N., Mengchen, W., Peigen, S., Vivek, B., and Cho, Y. H., "NetFPGA logic analyzer," *2nd North American NetFPGA developers workshop* (2010).

Paper 6: Tackling OpenFlow power hog in core networks with KeyFlow

S. Saldanã Cercós, R. E. Oliveira, R. Vitoi, M. Martinello, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, "Tackling OpenFlow power hog in core networks with KeyFlow," in *Electronic letters*, 2014.

Tackling OpenFlow power hog in core networks with KeyFlow

S. Saldaña Cercós, R.E. Oliveira, R.Vitoi, M. Martinello, M. R. N. Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy

We present a comprehensive data plane power consumption analysis of an OpenFlow 1.0 switch broken down into its design modules and propose KeyFlow as an alternative solution, since it eliminates flow table lookup reducing 53.7% of the overall power consumption.

Introduction: Transport software defined networking (T-SDN) aims at satisfying the requirements for scalability, flexibility, and reliability of the network of the future [1]. Although, T-SDN provides a flexible and programmable networking platform with the control plane detached from the data plane (which leaves the management intelligence in a centralized controller) [2], its impact on power consumption cannot be neglected upon equipment implementation. Network equipment power consumption accounts for 14% of the overall information and communications technology (ICT) sector [3]. Consequently, recent works have provided analyses on power consumption contributed by the optical and electrical layers [4], where more than 90% is attributed to the electrical layer [5].

On the electrical layer, T-SDN has emphasized the need of high performance switches, which makes ternary addressable content memories (TCAMs) an indispensable resource to provide such high throughput and packet processing rates [6]. Although TCAMs are pricey and power-demanding they are required for wildcard matching [7]. However, the power consumption problem is intensified in T-SDN where the number of fields that the wildcard match needs to handle is seven times higher than in traditional switching elements [6].

In this paper we expand the state-of-the-art beyond reported research by investigating for the first time the potential benefits for low power in addition to a low latency and low jitter approach: KeyFlow [8]. We present it as an alternative to compact TCAMs [6] for power saving since it is a solution for T-SDN, where table lookup is eliminated, leading to reduced core network fabric power consumption. In this work the implementation of a KeyFlow switch on a full line-rate 4-port 1 GbE NetFPGA-based is described. An analysis on power consumption broken down into static and dynamic power for a reference Open- Flow 1.0 switch is presented. The hardware design is based on the reference implementations from Stanford University [9]. To the best of our knowledge, this paper is the first to report on the OpenFlow power consumption for different NetFPGA resources, as well as for each design module. Moreover, a comparison between OpenFlow 1.0 and a novel KeyFlow switch is performed.

KeyFlow vs. OpenFlow: KeyFlow is an alternative to OpenFlow which aims to reduce core network routers' complexity by removing flow table lookup. By doing so, a relaxation of the forwarding engine is achieved, potentially reducing the power consumed by the TCAMs.

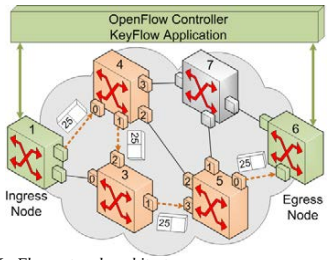


Fig. 1 *KeyFlow network architecture*

An example of a KeyFlow network architecture is depicted in Fig. 1, in which a flow from node 1 to 6 follows the end-to-end path through nodes m_i = (4; 3; 5) (dashed orange arrows) across the KeyFlow network, using the output ports p_i = (1; 1; 0), respectively. If there is no matching rule at the ingress node, the first packet of the flow is sent to the controller. The KeyFlow module in the controller computes a unique path ID (i.e. key). Once the key is computed (25 in this example), it needs to be assigned to the packet header by the ingress edge node. Therefore, the controller sends a rule to the edge nodes that i added in their flow table. Unlike traditional core nodes, KeyFlow nodes do not have a flow tables. Instead, an additional module, key processor, is added to the SDN implementation. When a packet with a given key in the header reaches the KeyFlow node, it performs a KeyFlow operation to determine the output port. This KeyFlow operation consists in computing the remainder of a division between the key and the node ID. Let $< X >_i$ be defined as the remainder from X/i. For the KeyFlow module, X is the key given by the controller and i is the node ID. Thereby, for the example presented in Fig. 1, the KeyFlow module performs the following operation to calculate the output port: $< 25 >_4 = 1$, $< 25 >_3 = 1$, and $< 25 >_5 = 0$, where 25 is the key, (4; 3; 5) are the nodes ID, and (1; 1; 0) are the respective output ports.

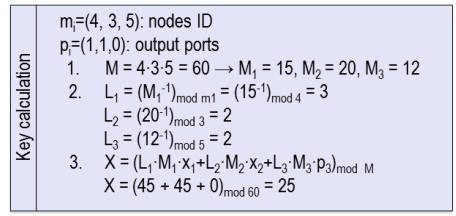


Fig. 2 *KeyFlow key calculation*

Fig. 2 illustrates an example for key calculation in the control plane based on the network presented in Fig. 1. Given a set of nodes m_i that defines a given path, the first parameter M_i is obtained by the division M/m_i, where M is the product of m_i. The second parameter, L_i, is the multiplicative inverse [10] of M_i (i.e. $M_i \cdot Li \equiv 1\ (mod\ m_i)$). Finally, the key, X, is obtained by the sum of the products $L_i \cdot M_i \cdot m_i\ (mod\ M)$ [8].

KeyFlow and OpenFlow implementation: There is a trend to build dedicated application-specific integrated circuits (ASICs) for SDN. Currently, FPGA is the industry sandbox to prototype them providing near-ASIC performance [7], [11]. Consequently, in this work, the KeyFlow switch has been implemented using a 4-port 1 GbE NetFPGA Xilinx Virtex-II Pro 50.

All reference designs with the official NetFPGA are based on the reference pipeline presented in [9]. The total power consumption results reported in this paper include MAC/CPU queues and user data path modules of the reference pipeline. However, this paper focuses on the optimization of the output port lookup from the user data path.

Fig. 3 presents the modular NetFPGA output port lookup pipeline structure for (a) a Hybrid switch which includes both KeyFlow and OpenFlow functionalities, and (b) a standalone KeyFlow switch. For the Hybrid switch operating as an OpenFlow switch from Fig. 3 (a), when a packet enters the output port lookup, the header parser gathers the relevant fields from the packet header and creates a flow to match against it. This flow is matched in two modules simultaneously: the exact match and the wildcard match.

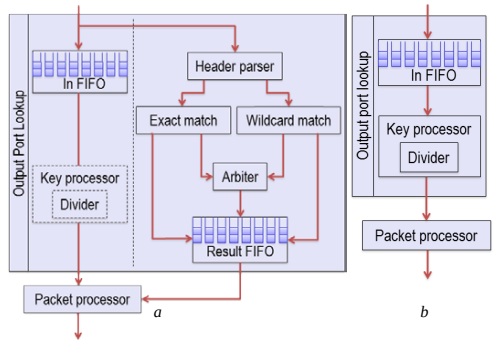


Fig. 3 *Modular NetFPGA pipeline structure for a) an Hybrid switch and b)a KeyFlow switch*

The exact match matches an exact flow entry using two hash functions and it uses the off-chip static random access memory (SRAM) to store the flow table, whereas the wildcard match is performed in the TCAMs build using the Xilinx SRL16e primitives [9]. The results from the matching modules are sent to the arbiter which determines the solution to be used (exact always wins over wildcard), and it writes the associated actions to the result first in, first out (FIFO) buffer.

The KeyFlow switch, however, reduces the complexity of the forwarding engine by substituting all the modules by a single module named key processor as shown in Fig. 3 (b), which performs the KeyFlow operation described earlier.

Methodology: To measure the power consumption of both OpenFlow 1.0 and KeyFlow switches, three steps are followed: First, a Perl-based testing infrastructure is used to perform simulation tests. The gate-level simulator, ModelSim, is used for the functional verification of the design, and to estimate frequencies of the clocks, activity rates of the logic resources, signals, input/outputs (IOs), the input clocks for the block random access memories (BRAMs), and digital clock managers (DCMs). Second, the design is synthesized using Xilinx ISE Virtex II 50 Pro NetFPGA. Third, the Xilinx Xpower Analyzer (XPA) tool from Xilinx ISE design software is used to calculate the power consumption. It uses the synthesized design from the second step and activity rates from the first step to increase the accuracy of the results.

Results: In order to present a comprehensive power consumption analysis, results include all evolving alternatives: from standalone OpenFlow switch to standalone KeyFlow switch, considering a hybrid solution as a transition. The results are detailed as: 1) Total; 2) Module-based; and 3) Resource-based power consumption for each switch. To avoid extra power consumption from buffering, tests are run generating a single 1500-byte packet. The flow table is kept fixed at 40 flows.

Table 1: Total power consumption [mW]

	Static	Dynamic	Total
OpenFlow	159	2514	2673
Hybrid	159	2384	2543
KeyFlow	159	1164	1323

Results presented in Table. 1 show that the static power consumption is 159 mW, which is independent on the switch implementation. However, in the dynamic power consumption, the Hybrid switch presents 6% net savings compared to the OpenFlow switch.

Fig. 4 *Module-based power consumption results*

Looking into the output port lookup module-based results presented in Fig. 4, the wildcard match is the power hog from the hybrid design, accounting for 78.75% of the output port lookup power consumption. By contrast, the key processor accounts only for 0.97%. These results explain the 6% net decrease since, simply by avoiding the table lookup process, the activity rates of the wildcard match (thereby the TCAMs) are reduced at almost no expenses from the addition of the key processor. Further improvements can be achieved by implementing a standalone KeyFlow switch.

Fig. 5 presents the total resource-based power consumption for the OpenFlow, Hybrid, and KeyFlow switches. By implementing the KeyFlow switch, signals and logic power consumption can be reduced 94.16% and 94.66%, respectively, compared to the Hybrid switch. These savings explain the overall dynamic power consumption of 1164 mW from Table.1, which represents a total 53.7% savings compared to an OpenFlow switch.

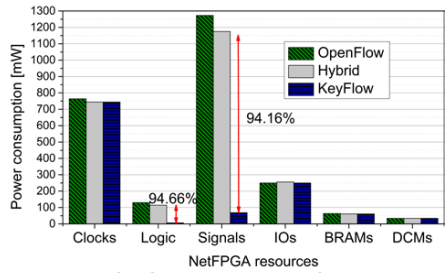


Fig. 5 *Resource-based power consumption results*

Conclusions: This paper presented for the first time a module- and resource-based power consumption analysis of a reference NetFPGA OpenFlow 1.0 switch. Additionally, a Hybrid and a KeyFlow switches were implemented and analysed. The results indicated that the major power consumption contributor is the wildcard match due to the power demanding TCAM. It accounts for 78.75% of the output port lookup Hybrid switch power consumption. A migration strategy towards the elimination of flow table lookup was suggested, since by implementing a Hybrid and a KeyFlow switch, 6% and up to 53.7% power savings were achieved, respectively. The results presented in the resource-based analysis showed the potential for further reducing power consumption by tackling the second major contributor: the clocks' frequencies. The development of power-aware OpenFlow-based SDN solutions on chip can benefit from the insides presented in this paper.

S. Saldaña Cercós, A. Manolova Fagertun, and I. Tafur Monroy (*DTU Fotonik, Technical University of Denmark, Kgs. Lyngby, Denmark*).

(ssce@fotonik.dtu.dk)

R.E. Oliveira, R.Vitoi, M. Martinello, and M. R. N. Ribeiro (*LABTEL-Electrical Engineering Department, Federal University of Espírito Santo, Vitória E.S. Brazil*).

References

1. P. Iovanna et al.: 'E2E traffic engineering routing for transport SDN', Proc. OFC, W1K.3 (2014).
2. D. Marconett et al.: 'Optical FlowBroker: load-balancing in Software-defined multi-domain optical networks,' Proc. OFC, W1K.3 (2014).
3. A. Vishmanath et al.: 'Adapting router buffers for energy efficiency,' CoNEXT, p.19, ACM (2011).
4. S. Saldaña et al.: 'Power-aware multi-layer translucent network,' Proc. OFC, W2A.53 (2014).
5. G. Rizzelli et al.: 'Energy efficient traffic-aware design of on-off multi-layer translucent optical networks,' computer Networks, Vol. 56, p.2443-2455 (2012).
6. K. Kannan et al.: 'Compact TCAM for power aware SDN,' ICDCN, Vol. 7730, p.429-443 (2013).
7. C. Zerbini et al.: 'Performance evaluation of packet classification on FPGA-based TCAM emulation architectures,' Globecom (2012).
8. M. Martinello et al.: 'KeyFlow: A prototype for evolving SDN toward core network fabric,' IEEE Network, March (2014).
9. J. Naous et al.: 'Implementing an OpenFlow switch on the NetFPGA platform.' Proc. ANCS (2008).
10. V. Rijmen: 'Efficient Implementation of the Rijndael S-box,' Katholieke Universiteit Leuven, Dept. ESAT. Belgium (2000).
11. W. Jiang et al.: 'Scalable ternary content addressable memory implementation using NetFPGAs.' ANCS (2013).

Paper 7: 100 Gbps IM/DD Links Using Quad-Polarization: Performance, Complexity, and Power Dissipation

S. Saldaña Cercós, M. Piels, J. Estaran, M. Usuga, E. Porto da Silva, A. Manolova Fagertun and I. Tafur Monroy, "100 Gbps IM/DD Links Using Quad-Polarization: Performance, Complexity, and Power Dissipation," in *Journal*, 2015

100 Gbps IM/DD links using quad-polarization: Performance, complexity, and power dissipation

S. Saldaña Cercós,[1*] M. Piels,[1] J. Estarán,[1] M. Usuga,[1] E. Porto da Silva,[1] A. Manolova Fagertun,[1] and I. Tafur Monroy[1]

[1] *Department of Photonics Engineering, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark*

[*] *ssce@dtu.fotonik.dk*

Abstract: A computational complexity, power consumption, and receiver sensitivity analysis for three different scenarios for short-range direct detection links is presented: 1) quad-polarization, 2) wavelength division multiplexing (WDM), and 3) parallel optics. Results show that the power consumption penalty associated to the quad-polarization digital signal processing (DSP) is negligibly small. However, the required analog to digital converters account for 47.6% of the total system power consumption. Transmission of 4x32 Gbps over 2 km standard single mode fiber is achieved with a receiver sensitivity of 4.4 dBm.

OCIS codes: (060.0060) Fiber optics and optical communications; (060.2360) Fiber optics links and subsystem; (060.4260) Multiplexing.

References and links

1. R. Vasan, "A venture perspective on cloud computing," in IEEE Computer Society **44**(3), 60-62 (2011).
2. T. Zuo, A. Tatarczak, M. Iglesias Olmedo, J. M. Estaran, J. B. Jensen, Q. Zhong, X. Xu, and I. Tafur Monroy, "O-band 400 Gbit/s client side optical transmission link," in Optical Fiber Communication Conference and Exhibition (OFC), (Optical Society of America 2014), paper M2E.4.
3. J. Wei, Q. Cheng, R. V. Penty, I. H. White, and D. G. Cunningham, "400 Gigabit Ethernet using advanced modulation formats: performance, complexity, and power dissipation," in IEEE Commun. Mag. **53**, 182-189 (2015).
4. H. Griesser, A. Dochhan, and J. P. Elbers, "DSP-enhanced transmission for 100G+ DWDM data center interconnects," in Signal processing in photonic communications, (Optical Society of America 2014), paper ST1D-4.
5. M. Iglesias Olmedo, T. Zuo, J. B. Jensen, Q. Zhong, X. Xu, S. Popov, and I. Tafur Monroy, "Multiband carrierless amplitude phase modulation for high capacity optical data links," in J. Lightw. Technol. **32**(4), 798-804 (2014).
6. D. Qian, N, Cviketic, J. Hu, and T. Wang, "108 Gb/s OFDMA-PON with polarization multiplexing and direct detection," J. Lightw. Technol. **28**(4), 484-493 (2010).
7. K. Kikuchi, "Electronic polarization-division demultiplexing based on digital signal processing in intensity-modulation direct-detection optical communication systems," in Optics Express **22**(2), 1971-11980 (2014).
8. K. Kikuchi, and S. Kawakami, "16-ary Stokes-vector modulation enabling DSP-based direct detection at 100 Gbit/s," in Optical Fiber Communication Conference and Exhibition, (Optical Society of America 2014), paper Th3K.6.
9. M. Morsy-Osman, M. Chagnon, M. Poulin, S. Lessard, and D. V. Plant, "1 224 Gb/s 10 km transmission of polarization division multiplexed PAM-4 signals using 1.3μm SiP intensity modulator and a direct-detection MIMO-based receiver," in European Conference and Exhibition on Optical Communication, (Optical Society of America, 2014), paper PD.4.4.

10. J. Estarán, M. Usuga, E. Porto da Silvia, M. Piels, M. Iglesias Olmedo, D. Zibar, and I. Tafur Monroy, "Quaternary polarization-multiplexed subsystem for high capacity IM/DD optical data links," in J. Lightw. Technol. **33**(7) 1408-1416 (2015).
11. C. Herard and A. Lacourt, "Three channel multiplexing using polarization of light," in Optics Communications **60**(1-2), 27-31v(1986).
12. P. Dong, C. Xie, and L. Buhl, "Monolithic polarization diversity coherent receiver based on 120-degree optical hybrids on silicon," in Optics Express **22**(2), 2119-2125 (2014).
13. L. Litwin, "Matched filtering and timing recovery in digital receivers," in RF design **24**(9), 32-49 (2001).
14. B. S. G. Pillai, B. Sedighi, K. Guan, N. P. Anthapadmanabhan, W. Shieh, K. J. Hinton, and R. S. Tucker, "End-to-end energy modeling and analysis of long-haul coherent transmission systems," in J. Lightw. Technol. **32**(18), 3093-3111 (2014).
15. Julius O. Smith, "Physical audio signal processing," Standford University, http://ccrma.standford.edu/ jos/pasp/, online book, date last viewed 10 (2010).
16. J. L. Wei, D. G. Cunningham, R. V. Penty, and I. H. White, "Study of 100 Gigabit Ethernet using carrierless amplitude/phase modulation and optical OFDM," in J. Lightw. Technol. **31**(9), 1367-1373 (2013).
17. Texas Instruments, "CMOS power consumption and cpd calculation," in SCAA035B (June 1997).
18. S. Savory, "Digital signal processing for coherent systems," in Optical Fiber Communication Conference and Exhibition, (Optical Society of America 2012).
19. K. Kikuchi, and S. Kawakami, "Multi-level signaling in the Stokes space and its application to large-capacity optical communications," in Optics Express **22**(7), 7374-7387 (2014).

1. Introduction

With the dot-com bubble in the late '90s came the boom of data centers [1]. However, with current traffic growth and the increase of the number of services that are powered by data centers, optical data links are suffering from bandwidth limitations. High capacity short range links are, therefore, of interest of research both in academia and at the industry level.

Enabling technologies to cope with high bitrate requirements include wavelength division multiplexing (WDM) and parallel optics; however, current traffic demands challenge the achievable net bitrate by WDM architectures and enhances the constraints introduced by the cabling overhead when using parallel optics. Thus, alternative solutions such as higher order modulation formats and space division multiplexing have been proposed. Higher spectral efficiency for data links has been explored employing 4 levels pulse amplitude modulation (4-PAM) [2], duo-binary modulation [3], and even more complex modulation formats such as discrete multitone (DMT) modulation [4] or carrierless amplitude phase (CAP) modulation [5].

This spectral efficiency increase comes at the expense of computational complexity, thus cost and power consumption [3]. Consequently, polarization multiplexing solutions had been proposed as an enabling technology for the deployment of high speed transmission intensity modulation direct detection (IM/DD) systems. In [6], self-coherent direct detection with a spectrally separated frequency reference is used for polarization-multiplexed transmission of an orthogonal frequency division multiplexed (OFDM) signal. The architecture also uses different frequency-orthogonal bands for each polarization to suppress crosstalk. Further improvements in terms of complexity are achieved by incoherent direct detection links [7]. Incoherent direct detection links are achieved by employing a Stokes receiver as reported in [8], and [9]. Based on the Stokes receiver structure from [7], in [10] a technique that allows for four independent data streams transmission on four different SOPs, quad-polarization, is presented.

The main goal of this work is to present a fair comparison between quad-polarization and the two competitive technologies: parallel optics, and WDM. In this paper an analysis of performance in terms of bit-error-rate (BER), digital complexity introduced by the digital signal processing (DSP) at the receiver, and power dissipation of the quad-polarization technique is reported as potential alternative to WDM and parallel optics. The results reported are based on 4-polarization multiplexing, 4-channel WDM, and 4-optical links at 32 Gbps transmission rate to achieve 100 Gbps net bitrate accounting for 28% forward error correction (FEC) overhead. Results are evaluated at the standard FEC limits (i.e., 7% and 20%), but since transmission was

achieved up to 32 Gbps per link, this allows for a higher FEC threshold and better sensitivity.

The rest of the paper is structured as follows: Section 2 provides a description of the quad-polarization concept, the requirements at the transmitter and receiver, receiver design, and required DSP. Section 3 determines the computational complexity of each of the blocks needed at the receiver. Power consumption results are reported in Section 4. Performance analysis in terms of bit error rate based on experimental results is presented in Section 5. Concluding remarks and discussion are summarized in Section 6.

2. Quaternary polarization multiplexing digital signal processing

This section presents a detailed description of: 1) the quad-polarization concept and the system testbed used, 2) the requirements at the transmitter that allow a quad-polarization system to operate without ambiguities at the receiver, 3) the receiver used and the motivation for using it, and 4) the DSP stages required to demodulate the data streams carried by four the states of polarization (SOPs).

The general idea of quad-polarization multiplexed systems consists of transmitting four data streams simultaneously over a single media (standard single mode fiber (SSMF)) using four different SOPs instead of using only the two conventional linear and orthogonal polarizations (X and Y). In previously demonstrated IM/DD systems a maximum of three linear SOPs were simultaneously transmitted and received without ambiguity [11], since only three independent polarization parameters were read given that phase was not recovered: the amplitude of the projections on the two perpendicular polarization planes, and the rotation angle. However, in [10] the optical phase is indirectly exploited to encode another independent stream in a circular SOP, thus increasing the multiplexing order by one.

Fig. 1 presents the block diagram of the testbed for the quad-polarization system. At the transmitter four distributed feedback lasers (DFBs) are used as light sources. The center frequency of each of the DFBs is spaced 100 GHz (denoted by DFB A= 193.2 THz through DFB D = 193.5 THz) to ensure incoherent power addition. With a monolithically integrated quad-polarization modulator, a single transmit laser could be used as the phase relationship between the four branches could easily be stabilized. In this case, the quad-polarization approach would increase system spectral efficiency. The DFBs feed four integrated Mach-Zehnder modulators (MZMs). The MZMs are driven by four independent 32 Gbd non-return-to-zero (NRZ) electrical signals with 3 V_{pp}. Each signal is derived from pseudorandom bit sequences (PRBS) of length $2^{15} - 1$, decorrelated with electrical delay lines. Four polarization controllers (PCs) are used at the output of the MZMs to obtained the four desired SOPs (i.e., X, Y, 45°, and

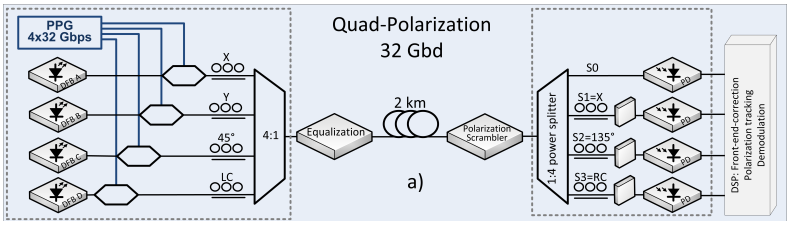


Fig. 1: Schematic of the experimental setup for a 4-SOP 128 Gbps transmission over 2 km standard singlemode fiber. DFB, distributed feedback laser; DSP, digital signal processing; PD, photodiode; PPG, pulse pattern generator; WDM, wavelength division multiplexing.

left-circular (LC)). Polarization multiplexing is performed using a standard 4:1 optical coupler. Before fiber transmission, a power equalization stage is required. The power equalization stage compensates for the varying MZM insertion losses at the transmitter. This is achieved using an erbium-doped fiber amplifier (EDFA) and a variable optical attenuator (VOA) (no net optical gain is provided through the power equalization stage), though it could also be achieved by manually adjusting the output powers of the lasers. Then the signal is launched into the 2 km SSMF for transmission.

In order to be able to demultiplex the four SOPs after transmission two key steps are taken: 1) compensation for the rotation of the SOPs when transmitted over SSMF, 2) demodulation in the digital domain. The Stokes analyzer and low-complexity DSP in the receivers are used to allow the transmission of 4-SOP optical signals.

Compensation of the Poincaré sphere rotation is achieved by using a polarization tracking algorithm and Stokes analyzer together with DSP at the receiver [7]. For quad-polarization, the work presented in [7] is extended by tracking three Stokes vectors instead of two. Thus, the rotation matrix of the fiber can be extracted, enabling signal demodulation of four independent SOPs. Details on the tracking algorithm are presented in Section 3. The polarization scrambler depicted in Fig. 1 at the end of the link is used for survey purposes. At this stage the polarization tracking algorithm performance is evaluated.

The Stokes analyzer used at the receiver is composed of four branches with four independent photodetectors which measure S0, S1, S2, and S3. Where S0 is the instantaneous total power, and S1, S2, and S3 are defined by Eqs. (1)–(3), respectively.

$$S1 = 2I_{X|Y} - S0 \tag{1}$$

$$S2 = 2I_{45|135} - S0 \tag{2}$$

$$S3 = 2I_{RC|LC} - S0 \tag{3}$$

At the first branch of the Stokes analyzer S0 is determined. An eye diagram of the quad-polarization signal received by the S0 photodiode is shown in Fig. 2(a). The remaining three branches use polarization controllers and polarizers to align to 1) X or Y defined as $I_{X|Y}$ to measure S1 , 2) 45° or 135° defined as $I_{45|135}$ to measure S2, and 3) right-circular or left-circular defined as $I_{RC|LC}$ to measure S3, respectively. Figures 2(b), (c), and (d) show the eye diagrams of the received signals by S1, S2, and S3 photodiodes, respectively. The signals received by these detectors change as a function of received state of polarization, but like the S0 eye, have multiple levels. After photo-detection the signal is sampled by a 80 GSa/s digital real time sampling scope (DSO) with 25 GHz bandwidth, and processed offline using DSP. The net insertion loss (IL) of this approach is between 1.2 and 3 dB for the system presented in this

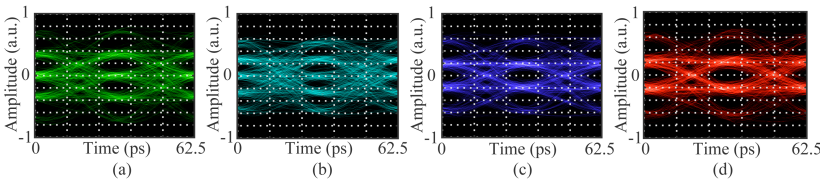


Fig. 2: Eye diagrams for the quad-polarization signal received by (a) the S0 photodiode, by (b) the S1 photodiode, by (c) the S2 photodiode, and by (d) the S3 photodiode.

work, depending on the fiber rotation. However, there can be a theoretically lossless system if an additional photodetector is used, [9], [12].

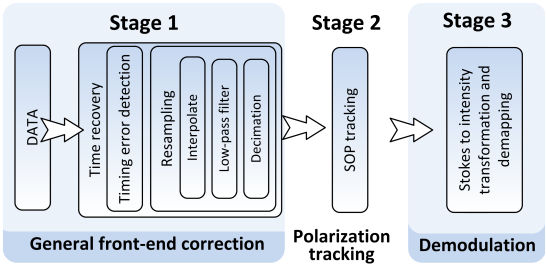


Fig. 3: DSP blocks for 4-SOP IM/DD. DSP, digital signal processing; IM/DD, intensity modulated/direct detection; SOP, state of polarization.

The second step consists of demodulation of the signal in the digital domain. Fig. 3 illustrates the three different stages required in the DSP at the receiver to successfully demodulate the data from the 4-SOP IM/DD system. In stage 1, general front-end corrections are performed. In stage 2 SOP tracking occurs, and finally in stage 3 the signal is demodulated.

Front-end correction includes error estimation (i.e., timing error correction) and resampling to the minimum number of integer samples per symbol. The resampling process is done in three steps: 1) interpolation, 2) a joint matched and anti-aliasing filtering, and 3) decimation. Note that for parallel optics and WDM systems a bit error rate tester (BERT) can be used at the receiver side eliminating the need for ADCs. In case of doing front-end corrections in the digital domain, then only one bit is required for the ADC reducing its power consumption by a factor four compared to the quad-polarization system.

After front-end corrections are preformed, the intensities are transformed to the Stokes parameters. In the Stokes space, propagation along the fiber can be described as a rotation. Tracking the received Stokes vectors amounts to tracking this rotation matrix.

Finally, at the third stage of the DSP, Stokes to intensity transformation and de-mapping are performed. Demodulation consists of a 4x4 multiple-input-multiple-output (MIMO) process. The main benefit in terms of computational load is that the mapping from the transmitter SOPs to the Stokes space and its inverse are known in advance as presented in Eq. (4)

$$\begin{bmatrix} I_X \\ I_Y \\ I_{45} \\ I_{RC} \end{bmatrix} = \begin{pmatrix} \\ M_{demux} \\ \\ \end{pmatrix} * \begin{bmatrix} S0 \\ S1 \\ S2 \\ S3 \end{bmatrix} \quad (4) \qquad \mathbf{M}_{demux} = \begin{pmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ 0 & 0 & 1 & 0 \\ 0.5 & -0.5 & -0.5 & -0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5)$$

where I_x, I_y, I_{45}, and I_{RC} are the transmitted intensities, and M_{demux} is the 4x4 demultiplexing matrix from Eq. (5).

Thus, the combined demapping and derotation process is described in Eq. (6)

$$\begin{bmatrix} I_x \\ I_y \\ I_{45} \\ I_{RC} \end{bmatrix} = \begin{pmatrix} \\ M_{demux} \\ \\ \end{pmatrix} * \begin{pmatrix} \\ M_{rot}^T \\ \\ \end{pmatrix} * \begin{bmatrix} S0 \\ S1 \\ S2 \\ S3 \end{bmatrix} \quad (6) \qquad \mathbf{M}_{rot} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & S_{11} & S_{12} & S_{13} \\ 0 & S_{21} & S_{22} & S_{23} \\ 0 & S_{31} & S_{32} & S_{33} \end{pmatrix} \quad (7)$$

where M_{rot} can be obtained from tracking three transmitted Stokes vectors according to Eq. (7).

$M_{rot}^{-1} = M_{rot}^{T}$ because M_{rot} is an orthogonal matrix. Consequently, no matrices need to be inverted in real time, reducing the digital processor requirements in terms of speed and power consumption.

3. Computational complexity analysis

In this section a computational complexity analysis for the aforementioned required DSP for the quad-polarization system is detailed. The computational complexity for each of the blocks described in Section 2 is measured by breaking each block down into number of operations (i.e., number of real multiplications and real additions) required for each of them.

For the general front-end correction, timing error detection is done using the Gardner algorithm. The error for the Gardner algorithm is computed using Eq. 8 [13] which accounts for one real addition and one real multiplication per symbol.

$$e_n = (y_n - y_{n-2}) \cdot y_{n-1} \tag{8}$$

Resampling is performed in three different steps. First, interpolation, then a low pass filter is applied, and finally, decimation is conducted. After decimation a matched filter is used. Table 1 summarizes the number of operations required for the timing error detector based on the work presented in [14], and for each of the three resampling steps.

Table 1: Number of real additions and real multiplications for Stage 1: front-end corrections

Action		*Num. of operations/s*	*Nomenclature*
Timing error detection	mult	$Baud \cdot Nspsym$	*Baud*, Baud-rate; *Nspsym*, Number of samples per symbol
	add	$Baud \cdot Nspsym$	
Interpolation	mult	$Baud \cdot Nspsym$	*Baud*, Baud-rate; *Nspsym*, Number of samples per symbol
	add	$2 \cdot Baud \cdot Nspsym$	
Low-pass filter	mult	$N_i \cdot Baud \cdot Nspsym$	N, Number of real taps for the filter i
	add	$N_i \cdot Baud \cdot Nspsym$	
Decimation	mult	$\frac{(D-1)}{D} \cdot Baud \cdot Nspsym$	*Baud*, Baud-rate; *Nspsym*, Number of samples per symbol; D, Decimation factor
	add	$\frac{1}{D} \cdot Baud \cdot Nspsym$	

Based on the work presented in [15], one-multiply form linear interpolation is assumed for resampling. One-multiply form linear interpolation has a computational complexity of one real multiplication and two real additions per sample of output. The number of samples of output is the baud-rate times the number of samples per symbol. After interpolation a an anti-aliasing low pass filter is used to suppress the "ghost frequencies". The computational complexity depends on the number of filter taps (N_1), which corresponds to the filter length (i.e., order) minus

one filter tap. One real multiplication and one real addition are used to implement a real filter tap. Notice that no complex taps are used for filtering, thus reducing the computational complexity by at least a factor three. Decimation is the last step for resampling. The computational complexity depends on the decimation factor D, which determines how many data points are eliminated. A decimation factor of three indicates that two out of three samples is eliminated. For the work presented here, one real multiplication for eliminated data point (i.e., $\frac{(D-1)}{D}$), and one addition per stored sample (i.e., $\frac{1}{D}$) is accounted for. The decimation factor used for the quad-polarization scenario was the sampling rate. Analogous to the anti-aliasing filter, the matched filter computational complexity is proportional to the number of the filter real taps (N_2).

After the general front-end correction, SOP tracking is performed. Fig. 4 illustrates the required sub-blocks for the SOP tracking algorithm.

The DSP here described and the tracking algorithm are based on the work presented in [7]. The algorithm described in [7] is implemented for two SOPs (the conventional orthogonal X and Y polarizations). This same algorithm is extended for quad-polarization by carefully defining the transmitted Jones vectors, as well as the receiver design and DSP described in Section 2.

The first step in stage 2 for the polarization tracking is to generate the Stokes vectors. For that Eqs. (1)–(3) are used. S_0 from the Stokes vectors indicates the instantaneous total intensity. As presented in Table 2, generating the Stokes vectors requires three multiplications and three additions (one for each branch of the Stokes analyzer) per sample. Note that the number of samples per symbol is reduced to one after decimation in the front-end corrections.

Once the signal has been transformed to the Stokes environment, an intensity discriminator is required. It discriminates zero-power symbols in order to ensure that no mathematical indeterminates will appear during the normalization process. In order to define the zero-power symbols, a threshold S_{th} is established so that all $S_0(n)$ symbols with intensity below the threshold level are de-mapped and removed from the four Stokes sequences. This step requires one addition per symbol.

Then, normalization of the Stokes vectors by $S_0(n)$ is performed. This operation maximizes the polarized components power. This process requires one multiplication per symbol.

Next, an iterative process is performed where an amplitude discrimination and a reference update take place. This process tracks the Stokes vectors associated with two or three transmitted signals. This tracking algorithm is based on the work presented in [7] and it is performed in four steps: 1) define two or three reference unit Stokes vectors $v(n)$, 2) determine the normal-

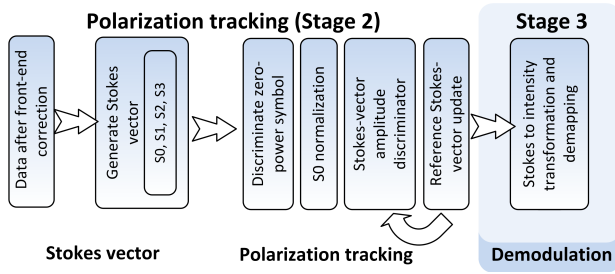


Fig. 4: DSP blocks for stage 2: SOP tracking

Table 2: Number of real additions and real multiplications for Stage 2: SOP tracking. *Baud*, number of samples after decimation (i.e., baudrate); L_{th}, number of threshold updates; δ, number of reference updates.

Action		*Num. of operations/s*	*Mathematical expr.*
Generate Stokes vectors	mult	$3 \cdot Baud$	Equations (1)–(3)
	add	$3 \cdot Baud$	
Intensity discriminator	mult	0	$S_0(n) - S_{th} < 0$
	add	$Baud$	
Normalization	mult	$Baud$	$\frac{S(n)}{S_0(n)}$
	add	0	
Amplitude discriminator	mult	$2 \cdot Baud$	$u(n) - u_{th_i} < 0$ $(for\ i = 1,2,3)$ and Eq. (9)
	add	$2 \cdot Baud$	
Reference updater	mult	$3 \cdot Baud \cdot \delta$	Equation (10)
	add	$4 \cdot Baud \cdot \delta$	
Threshold updater	mult	$2 \cdot Baud \cdot L_{th}$	$L_{th} = \frac{1}{100}$
	add	$3 \cdot Baud \cdot L_{th}$	

ized Stokes vector along the direction of the reference Stokes vector as presented in Eq. (9),

$$u(n) = \left[\frac{S(n)}{S_0(n)}\right] \cdot v(n) \tag{9}$$

3) define two or three thresholds u_{th_1}, and u_{th_2} which are used to determine which tributary the measured sample belongs to (i.e., which SOP is used at the transmitter side). If $u(n) \geq u_{th_1}$ then the measured sample belongs to the first tributary, if $u(n) \geq u_{th_2}$ then it belongs to the second tributary, etc., and 4) update the reference Stokes vector based on which tributary the measured sampled belongs to. Equation (10) indicates the reference Stokes vector update when the measured sample belongs to the tributary i, where μ is the step-size parameter.

$$v_i(n+1) = \frac{v(n) + \mu\left[\frac{S(n)}{S_0(n)} - v(n)\right]}{\left|\left|v(n) + \mu\left[\frac{S(n)}{S_0(n)} - v(n)\right]\right|\right|} \tag{10}$$

The case where $u(n) < u_{th_1}$ and $u(n) < u_{th_2}$ indicates that the transmitted SOP is not one of the SOPs being tracked and thus, $v(n+1)$ is not updated. The number operations needed for the reference updater is based on how often the Stokes vector is updated (δ from Table 2). Based on Eq. 10 the number of real additions and real multiplications is 4 and 3, respectively.

If the three tracked vectors are the signals transmitted in X, 45, and right-circular polarizations, the rotation matrix of the fiber can then be determined from the three tracked vectors according to the Eq. (7). The third vector can either be tracked using the same algorithm as v_1 and v_2, or from their cross product (i.e., $v_3 = v_1 \times v_2$).

The performance of both options is similar. The quad-polarization constellation consists of 16 distinct symbols. For each tracked vector, only one of these is used. Assuming that the data

is independently and identically distributed, this corresponds to an update parameter δ of 2/16 for tracking two vectors and 3/16 for tracking three vectors. Thus, from the two possible flows aforementioned, in case of obtaining v_3 from cross-product, the parameter δ is $2/16$, and 6 additional multiplications and 3 additions are required for the cross product. In case of tracking the three vectors v_1, v_2, and v_3, then the parameter δ is 3/16. For the calculations presented in Section 4 the second flow is used (i.e., tracking three vectors). The reader is referred to [7] for further details on the mathematical formulation of the tracking algorithm.

Additionally, the two (three) thresholds, u_{th_1} and u_{th_2} (u_{th_3}) need to be updated occasionally to assure convergence. Threshold update requires three additions (to find which signal is the strongest among the ones received by S1, S2, and S3), and two multiplications (one to normalize, and one to multiply by a tuning factor). Threshold update occurs 1/100 of the time as presented in Table 2. This value has been chosen empirically.

Finally, stage 3 deals with demodulation. For this Stokes to intensity transformation and demapping are performed. These two processes can be performed with a single matrix multiplication times the received vector $u(n)$ as presented in Eq. (6)[10]. Table 3 summarizes the number of operations required for demodulation, which enclose 4x4x4 multiplications and 3x4x4 additions for demultiplexing.

Table 3: Number of real additions and real multiplications for Stage 3: demodulation.

Action		*Num. of operations*	*Mathematical expr.*
Demultiplexing	mult	$64 \cdot Baud$	Eq. (6)
	add	$48 \cdot Baud$	

4. Power consumption

Power consumption is currently an important criterion to evaluate the feasibility of 100 Gbps IM/DD links [16]. Consequently, this section provides first, a power consumption analysis based on the number of operations for each DSP block. Second, in order to define the power consumption for each of the scenarios, the system layouts for quad-polarization, WDM, and parallel optics are described, providing with an overview of all optical and electronic components needed for 4x32 Gbps transmission. Finally, numerical results are reported.

4.1. Quad-polarization DSP energy consumption

The number of real additions and multiplications depends on the specific DSP implementation. The authors would like to emphasize that DSP has not been optimized for power consumption and consequently, results here present serve as an overview and estimate rather than exact and optimized estimates. The total power consumption also depends on the specific application-integrated circuit (ASIC) used. The numerical results hereby presented are based on a commercially available ASIC to provide a numerical approximation rather than an exact estimate since the ASIC power consumption can be reduced by modifying the frequency of operation or the supply voltage of the CMOS technology [17]. An ASIC based on 90 nm CMOS process technology is considered. Nortel provides detailed information for one of their ASIC designs which is suitable for 32 Gbd and consumes 1.5 pJ and 0.5 pJ per real multiplication and real addition, respectively [18]. These values have been used for power consumption evaluation of the quad-polarization DSP. The energy consumption associated to the DSP is described in Eq. (11).

$$E_{DSP} = E_{front-end} + E_{track} + E_{dem} \tag{11}$$

Where $E_{front-end}$, E_{track}, and E_{dem} are the energy associated to the front-end corrections, SOP tracking, and demodulation, respectively. The energy for each DSP block has been calculated by accounting for the number of multiplications and additions described in Section 3 and the average power consumption per real multiplication and additional of the aforementioned ASIC.

Table 4: Modelling parameters - power consumption

Parameter	**Definition**	**Value**
$Baud$	Baudrate	$32 \cdot 10^9$
$Nspsym$	Number of samples per symbol	16
N_1	Anti-aliasing filter taps	640
N_2	Matched filter taps	2
D	Decimation factor	$80 \cdot 10^9$
δ	Number of reference updates	$3/16$
L_{th}	Number of threshold updates	$1/100$
F_A	ADC figure of merit	$2.5 \cdot 10^{-12}$ J/conv-step
n_{adc}	ADC resolution	4 bits
DFB	Distributed Feedback laser	2.5 W
MZM	Mach-Zehnder modulator	832 mW
PD	Photo-detector	25 mW

Table 4 summarizes the parameters used and their values for the DSP power consumption calculation as well as the specifications of the opto-electronic components. The values are based on the experimental demonstration reported in [10]. Results on DSP power consumption are presented in Table 5. For WDM and parallel optics DSP it is assumed only front-end corrections are required. Implementations of both technologies that do not use DSP are possible. Power consumption analysis for both scenarios is provided in Section 4.3. Quad-polarization DSP power consumption is divided in two sections: 1) the initialization process for the SOP tracking algorithm (i.e., SOP tracking), and 2) the demodulation process. The power consumption associated to the polarization tracking algorithm depends on how fast the algorithm converges (i.e., what is the number of iterations needed in order to find an accurate estimate for $v(n)$). In [7] a 1000 sample period for the tracking algorithm to converge is presented, and the maximum convergence depth has been reported to be a 2000 sample period [10]. However, for the results presented in Table 5, the use of a training sequence is assumed, reducing the number of iterations to one [19]. For the quad-polarization DSP two additional power consumption contributions have to be considered: the reference update, and the threshold(s) update. These two processes occur 3/16 symbols (when tracking 3 vectors), and 1/100 symbols, respectively. For stage 3 (i.e., demodulation) of the quad-polarization scenario, power consumption is split in two: the demodulation process based on Eq. (6) and the power consumption contribution of the analog to digital converters (ADCs). The power consumption of the ADC is calculated based on Eq. (12) [14].

$$E_{ADC} = 4F_A n_{adc} F_s \tag{12}$$

Where F_A is a figure of merit, n_{adc} is the nominal ADC resolution (ADC resolution refers to the physical ADC resolution, and it is considered to be two bits higher than the effective number of bits [14]), and F_s is the sampling rate. As presented in Table 5 the majority of the power consumption, being 658.7 mW, for DSP is associated to the front-end corrections, which are common for the three scenarios. The additional DSP needed for SOP tracking and demodulation in the quad-polarization scenario is negligible (0.4 and 3.8 mW, respectively). However, as many ADCs as SOPs are required, which are one of the most power demanding elements. Each ADC consumes 3.2 W, thus, consuming 12.8 W for the quad-polarization case. One fourth of this power consumption is accounted for ADCs for parallel optics and WDM.

Table 5: Power consumption

	Stage	Power consumption
WDM	Front-end correction	658.7 mW
Parallel optics		
Quad-polarization	SOP tracking	0.4 mW
	Demodulation	3.8 mW
	ADC	12.8 W

4.2. System layout

The power consumption analysis of the three scenarios includes the contribution of the required optical and electronic components for each scenario. Thus, in this section the testbed used for the WDM system and parallel optics are described and compared to the quad-polarization testbed described in Section 2.

Fig. 5 presents the block diagram for (a) WDM and (b) parallel optics. The first part of the transmitter side, which was described in Section 2 is common for the three scenarios. For the WDM subsystem no power equalization stage is required before transmission. However, it is maintained in the setup in order to assure the same transmitted signal-to-noise ratio (SNR) as for the quad-polarization case. The four WDM signals are multiplexed with a standard 4:1 optical coupler.

After transmission and wavelength demultiplexing in an arrayed waveguide grating, the NRZ signals are detected by four PDs and sampled by the DSO described above. The stored signals are processed offline using DSP to perform front-end correction, demodulation, and BER evaluation. In the parallel optics scenario the signals are sent independently through four independent fibers. The experimental setup was implemented using a single SSMF and a switch to take the measurements from channel A to channel D as shown in Fig. 5(b) to assure the same transmission link performance for all the channels. Analogous to the WDM subsystem, the power equalization stage and the PCs at the receiver side are not required in the set-up, although maintained to assure the same transmitted SNR and photodiode insertion loss as for the quad-polarization system in order to offer a fair comparison between the three scenarios.

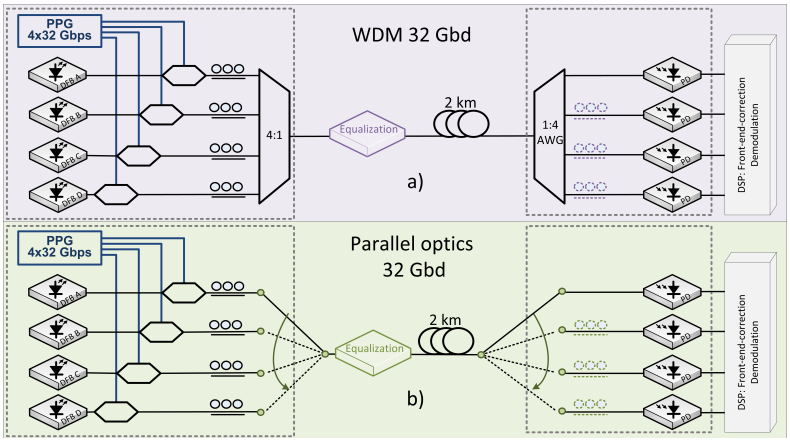


Fig. 5: Schematic of the experimental setup for (a) 4 channel WDM, (b) 4-lane parallel optics 128 Gbps transmission over 2 km standard single mode fiber. DFB, Distributed feedback laser; DSP, Digital signal processing; PD, Photodiode; PPG, pulse pattern generator; WDM, Wavelength division multiplexing

4.3. Power consumption numerical results

The power consumption from the electrical signal generation, light sources (i.e., DFBs), modulators (i.e., MZMs), and photo-detectors (i.e., PDs) is considered a base contribution common in the three scenarios. The power consumption contribution of the power equalization stage for the quad-polarization scenario is not accounted for in the overall power consumption since the equalizer could be implemented as a passive component. The remaining optical and electronic components for both WDM and parallel optics are passive optics, and thus their contributions to the overall power consumption are negligible. However, it should be noted that for some WDM channel spacings and deployment conditions active temperature control is needed in order to maintain alignment between the transmitter and wavelength demultiplexer. The DSP developed

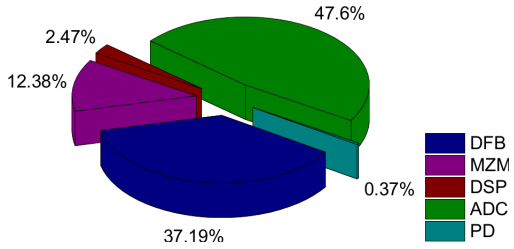


Fig. 6: Quad-polarization components power consumption. ADC, analog to digital converter; DFB, distributed feedback laser; DSP, digital signal processing; MZM, Mach-Zehnder modulator; PD, Photo-detector.

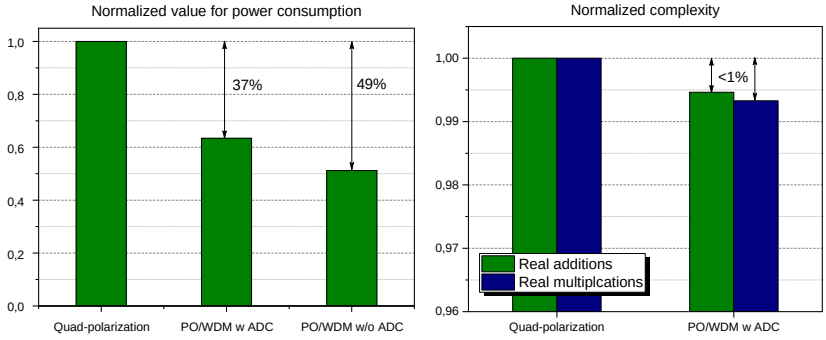


Fig. 7: Normalized values for (a) power consumption and (b) DSP complexity for the quad-polarization and parallel optics and WDM with and without ADC. ADC, analog to digital converter; PO, parallel optics; WDM, wavelength division multiplexing

in Section 2 is used to analyze the receiver energy consumption, where only stage 1 and a simple hard decision stage are used for both WDM and parallel optics. Additional polarization tracking and demapping stages are accounted for in the quad-polarization scenario.

For the base power consumption contribution, 2.5 W, 832 mW, and 25 mW are considered for each light source, modulator, and photo-detector, respectively [14]. Fig. 6 presents the contributions of each of the components for the quad-polarization scenario. The additional DSP for polarization tracking accounts for only 2.47% of the total power consumption due to its low complexity, however power consumption is increased by 47.6% due to the use of power-hungry ADCs.

The power consumption and complexity values normalized to the quad-polarization system are presented in Fig. 7(a) and Fig. 7(b), respectively. Results show a power consumption penalty of 37% for the quad-polarization compared to parallel optics and WDM systems due to the more demanding ADCs contribution. Up to 49% increased power consumption is observed when no ADC is used for parallel optics and WDM. However, the increased complexity associated to the DSP remains below 1%.

5. Performance

In this section a transmission performance in terms of bit-error-rate (BER) comparison between quad-polarization, 4-channel WDM system, and parallel optics is presented. The configuration evaluation has been assessed for 4x32 Gbps back-to-back (B2B) and 2 km of fiber type G.652 standard single mode fiber (SSMF) (16.5 ps/nm· km chromatic dispersion, 0.2 dB/km attenuation) transmission. The BER results presented account for the entire system, and the received power in all cases corresponds to the total received power, not the received power per photodetector. In all systems some variation in sensitivity between channels is observed.

5.1. *Results*

For the quad-polarization case, the BER performance of the system for perfect polarization rotation compensation post-convergence is plotted in Fig. 8 with black square symbols for the B2B case and red circle symbols for the 2 km transmission case. FEC limits for both 7% and 20% are illustrated for reference purposes. BER below FEC limits is obtained both for B2B and 2 km transmission. The receiver sensitivity for 100 Gbps net bitrate assuming 7% FEC overhead is 4.4 dBm. A penalty of 0.5 dB is observed after transmission. The B2B receiver sensitivity for 20% FEC overhead is 3.7 dBm with 0.5 dB measured penalty after 2 km SSMF transmission. For the WDM system, four 32 Gbd NRZ data signals for each WDM channel were successfully recovered after 2 km transmission. In Fig. 9 the BER for B2B (red circles)

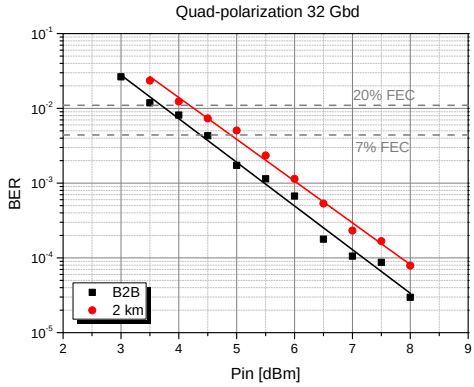


Fig. 8: BER sensitivity to PD input power for 32 Gbd quad-polarization for B2B (black, square) and 2 km SSFM transmission (red, circle).

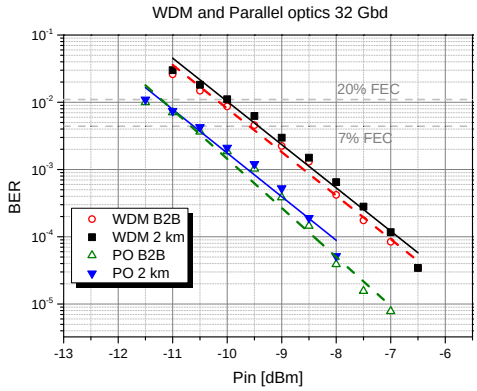


Fig. 9: BER sensitivity to PD input power for 32 Gbd for WDM for B2B (red, circle) and 2 km SSFM transmission (black, square), and for Parallel optics for B2B (green, triangle) and 2 km SSMF transmission (blue, triangle).

and after transmission (black squares) is computed as a function of the input power into the receiver for the entire WDM system. The receiver sensitivity for 100 Gbps net bitrate is -9.5 dBm and -10.2 dBm with no transmission penalty measured assuming 7% FEC and 20% FEC overhead, respectively. Finally, for the parallel optics scenario, results presented in Fig. 9 show a receiver sensitivity for 100 Gbps net bitrate of -10.5 dBm and -11.3 dBm for 7% and 20% FEC overhead, respectively. The 1dB improvement with respect to the WDM case is primarily due to the insertion loss of the AWG.

6. Conclusion

This work presents a comparison between quad-polarization, 4-channel WDM, and 4-line parallel optics in terms of computation complexity, power consumption, and system receiver sensitivity. The study presented in this paper includes a detailed analysis on the additional DSP required for the quad-polarization scenario at the receiver, and its power consumption contribution based on the number of operations needed for each of the DSP blocks. Quad-polarization for direct-detection optical subsystems allows the transmission of four parallel data streams using for different states of polarization. This approach has the potential to increase the capacity per channel. However, an increase in capacity comes at expenses of both power consumption and receiver sensitivity. Results show that quad-polarization DSP complexity is very low presenting only 2.47% additional power consumption compared to WDM and parallel optics. However, ADCs are required which account for 47.6% of the system total power consumption.

In terms of BER parallel optics is the system that provides lower receiver sensitivity, -10.5 dBm for 7% overhead FEC with no transmission penalty. For the WDM system 1 dBm penalty is observed, though reducing the footprint (i.e., cabling) by a factor of four. For the quad-polarization system 4.4 dBm receiver sensitivity is obtained, with 0.5 dBm penalty after transmission. This sensitivity penalty is expected since it represents the comparison between a single level approach versus a multi-level alternative. Better sensitivity could be achieved with dual polarization solutions.

Quad-polarization combined with WDM, and/or parallel optics offers an alternative solution for 400 Gbps low complexity short-reach optical transmission systems since it enables bitrate quadrupling for each laser-photodiode-ADC lane but comes with the costs in performance typically associated with multilevel modulation formats. Additionally, quad-polarization can be considered as a competitor solution to recent standarization activities such IEEE 802.3 standard, 400 Gigabit Taskforce adopted PAM 4 for 100 Gbps per wavelength. Relative performance, complexity, and power dissipation of the quad-polarization scheme using 100 Gbps PAM 4 as a reference for comparison would be an interesting topic for future research.

Bibliography

[1] ITU, "The world in 2014: ICT facts and figures", *in ICT statistics* (2014).

[2] [online available:] www.internetworldstats.com

[3] Cisco inc, "Cisco Visual Networking Index: Forecast and methodology, 2013-2018" June (2014).

[4] The Climate Group, "Smart 2020: Enabling the Low Carbon Economy in the Information Age", *in Report for Global e-Sustainability Initiative*, GeSI, 1-86, (2008).

[5] C. Lee, K. Yoontae, and J. K. Rhee, "Green IP over WDM network design considering energy-load proportionality," *in International conference on computing, networking, and communications* (2012).

[6] NTRC, "A review of the energy tax on electric power consumption," *in NTRC tax research journal*, **23**(5), 11-41 (2011).

[7] J. Baliga, M. Feng, R. Ayre, and R. S. Tucker, "Power consumption and energy efficient in the internet," *in IEEE Network*, **25**(2), 6-12 (2011).

[8] M. Ajmone, C. Guerrero, S. Buzzi, F. Idzikowski, L. Chiaraviglio, M. Meo, Y. Ye, and J. López Vizcaíno, "TREND: Toward real energy-efficient network design," *in TREND project*, 1-6 (2012).

[9] F. J. Effenberger, J. I. Kani, and Y. Maeda, "Standardization trends and prospective views on the next generation of broadband optical access systems," *in IEEE Journal on Selected Areas in Communications*, **28**(6), 773-780 (2010).

[10] K.Papaefthimiou, Y. Tefera, D. Mihaylov, J. M. Guitiérrez, and

M. Jensen, "Algorithmic PON/P2P FTTH access network design for CAPEX minimization," *in Telecommunication forum (TELFOR)*, 168-171 (2013).

[11] Alcatel-Lucent Bell Labs research, "Upscaling the metro, deploying the 7950 XRS in the metro core," *Application note* (2015).

[12] M. Henriques, P. Pinho, and A. Teixeira, "Energy-aware routing and grooming for IP transport over WDM MLR networks," *in Telecommunication network strategy and planning symposium (Networks)*, 1-7 (2014).

[13] A. N. Patel, C. Gao, J. P. Jue, X. Wang, Q. Zhang, P. Palacharla, and T. Naito, "Traffic grooming and regenerator placement in impairment-aware optical WDM networks," *in ONDM* (2010).

[14] J. López Y. Yabin, F. Jiménez, and P. M. Krummrich, "Amplifier placements optimization for enhanced energy efficiency in optical transport networks," *in Communications ICC* 3700-3706 (IEEE International Conference 2014).

[15] M. Martinello, M. R. N. Ribeiro, R. E. Z. de Oliveira, and R. de Angelis Vitoi, "KeyFlow: A prototype for evolving SDN toward core network fabrics," *IEEE Network*, **28**, 12-19 (2014).

[16] J. G. Koomey, "Growth in data center electricity use 2005 to 2010," *A report by Analytical Press, completed at the request of New York Times* (2011).

[17] J. Estarán, M. Usuga, E. Porto da Silvia, M. Piels, M. Iglesias Olmedo, D. Zibar, and I. Tafur Monroy, "Quaternary Polarization-Multiplexed Subsytem for High Capacity IM/DD Optical Data Links," *in Journal of Lightwave Technology*, **33**(7) 1408-1416 (2015).

[18] S. Savory, "Digital signal processing for coherent systems," *in Optical Fiber Communication Conference and Exhibition*, (Optical Society of Ameirca 2012).

[19] Tektronix, "Techiniques for extending real-time oscilloscope badwidth," *in White paper tektronix* (2014).

[20] S. Saldaña Cercós, L. C. Resendo, M. R. N.Ribeiro, A. Manolova Fagertun, and I. Tafur Monroy, "Power-aware multi-layer translucent network design: an integrated OPEX/CAPEX analysis,"

in Optical Fiber Communication Conference and Exhibition, (Optical Society of Ameirca 2014), paper W2A.53.

[21] W. Hou, L. Guo, X. Wei, and X. Gong, "Multi-granularity and robust grooming in power- and port-cost-efficient IP over WDM networks," *in Computer networks*, **56**(10), 2383-2399 (2012)

[22] S. Spolitis, C. wagner, J. J. Vegas Olmos, V. Bobrovs, G. Ivanovs, and I. Tafur Monroy, "Experimental demonstration of a scalable sliceable transceiver for optical access network," *in Asia Communications and Photonics Conference* (2014).

[23] S. Spolitis, J. J. Vegas Olmos, V. Bobrovs, G. Ivanovs, and I. Tafur Monroy, "A novel approach for transmission of 56 Gb/s NRZ signal in access network using spectrum slicing technique," *in Asia Communications and Photonics Conference* (2013).

[24] I. Tsirilakis, C. Mas, and I. Tomkos, "Cost comparison of IP/WDM vs. IP/OTN for European backbone networks," *in Proc. International conference of transparent optical networks*, **2**, 46-49 (2005).

[25] P. Mahajan, "The economics of using IP over OTN in core networks," *in Proc. World telecommunications congress*, 1-3 (2014).

[26] E. Palkopulou, D. A. Schupke, and T. Bauschert, "Energy efficiency and CAPEX minimization for backbone network planning: Is there a tradeoff?," *in Advanced networks and telecommunication systems*, 1-3 (2009).

[27] I. Stiakogiannis, E. Palkopoulou, and I. Tomkos, "On the cost efficieny of flexible optical networking compared to conventional SLR/MLR WDM networks," *in Proc. International conference on transparent optical networks* (2013).

[28] G. Rizzelli, A. Morea, M. Tornatore, and O. Rival, "Energy efficient traffic aware design of on-off multi-layer translucent optical networks," *in IEEE Computer networks*, **56**(10), 2443-2455 (2012).

[29] Cisco, "Cisco visual networking index: Forecast and methodology, 2012-2017," *in Cisco white paper* (2013).

[30] L. C. Resendo, M. R. N. Ribeiro, and J. Jo Pires, "Optimal multilayer grooming-oriented design for inter-ring traffic protection in DNI multi-

ring WDM networks," *in Journal of optical networking*, **7**(6), 533-549 (2008).

[31] M. Elbeshti, M. Dixon, and T. Koziniec, "Cost effective RISC core supporting the large sending offload," *in Communications and information technologies*, 844-849 (2012).

[32] F. Musumeci, D. Siracusa, G. Rizzelli, M. Tornatore, R. Fiandra, and A. Pattavina, "On the energy consumption of IP-over-WDM architectures," *in Proc. International conference on communications* (2012)

[33] F. Idzikowski, S. Orlowski, H. Woesner, and A. Wolisz, "Saving energy in IP-over-WDM networks for switching off line cards in low-demand scenarios," *in Proc. Optical networks design and modelling* (2011).

[34] Y. Zhang, M. Tornatore, P. Chowdhury, and B. Murkherjee, "Energy optimization in IP-over-WDM networks," *in Optical switching networks*, **8**(3), 177-180 (2011).

[35] A. Morea, O. Rival, N. Brochier, and E. L. Rouzic, "Datarate adaptation for night-time energy savings in core networks," *Journal of Lightwave Technology*, **31**(5), 779-785 (2013).

[36] W. V. Heddegehem, F. Idzikowski, W. Vereecken, D. Colle, M. Pickavet, and P. Demeester, "Power consumption modeling in optical multilayer networks," *in Photonics network communications*, **23** (2012).

[37] A. Klekamp U. Gebhard, and F. Ilchmann, *Energy and cost efficiency of adaptive and mixed-line-rate IP over DWDM networks,* in Journal of lightwave techonology, vol. 30, n.2, pp. 215-221, 2012.

[38] W. V. Heddegehem, F. Idzikowski, E. L. Rouzic, J. Y. Mazeas, H. Poignant, S. Salaun, B. Lanoo, and D. Colle, "Evaluation of power rating of core network equipment in practical deployments," *in online conference on green communications (GreenCom)*, 123-132 (2012).

[39] G. Rizzelli, A. Morea, M. Tornatore, and A. Pattavina, "Reach-related energy consumption in IP-over-WDM 100 G translucent networks," *in Journal of lightwave technology*, **31**(11) (2013).

[40] Agrawal, Godvin P., "Fiber-optic communication systems," *in Wiley* (2010).

[41] S. Aleksic, "Energy and entropy flow in erbium-doped fiber amplifiers:

a thermodynamic approach," *in Journal of lightwave technology* **30**, 2832-2838 (2012).

[42] B. Heller, R. Sherwood, and N. McKeown, "The controller placement problem," *in HotSDN'12*, New York, NY, USA (2012).

[43] D. Hock, M. Hartmann, S. Gebert, M. Jarschel, T. Zinner, and P. Tran Gia, "Pareto-optimal resilient controller placement in SDN-based core networks," *in Proc. International teletraffic congress* (2013).

[44] L. Liu, T. Tsuritani, I. Morita, R. Casellas, R. Martínez, and R. Muñoz, "Control plane techniques for elastic optical networks: GMPLS/PCE vs OpenFlow," *in IEEE Globecom Workshops* (2012).

[45] J. Naous, D. Erickson, G. A. Covington, G. Appenzeller, and N. McKeon,"Implementing an OpenFlow switch on the NetFPGA platform," *in Proc. of Architectures for networking and communication systems* (2008).

[46] C. A. Zerbini, and J. M. Finochietto, "Performance evaluation of packet classification on FPGA-based TCAM emulation architectures," *in GlobeCom* (2012).

[47] A. Mejia, J. Flich, J. Duato, S. A. Reinemo, and T. Skeie. "Segment-based routing: an efficient fault-tolerant routing algorithm for meshes and tori," *in Parallel and distributed processing symposium* (2006).

[48] K. Hasharoni, S. Benjamin, A. Geron, S. Stepanov, G. Katz, I. Epstein, N. Margalit, and M. Mesh, "A 1.3 Tb/s parallel optics VCSEL link," *in SPIE OPTO* (International Society for Optics and Photonics, 2014).

[49] T. Zuo, A. Tatarczak, M. Iglesias Olmedo, J. M. Estaran, J. B. Jensen, Q. Zhong, X. Xu, and I. Tafur Monroy, "O-band 400 Gbit/s client side optical transmission link," *in Optical Fiber Communication Conference and Exhibition*, (Optical Society of Ameirca 2014), paper M2E.4.

[50] J. Wei, Q. Cheng, R. V. Penty, I. H. White, and D. G. Cunningham, "400 Gigabit Ethernet using advanced modulation formats: Performance, complexity, and power dissipation," *in IEEE Communications Magazine*, **53**, 182-189 (2015).

[51] H. Griesser, A. Dochhan, and J. P. Elbers, "DSP-enhanced transmission for 100G+ DWDM data center interconnects," *in Signal process-*

ing in photonic communications, (Optical Society of America 2014), paper ST1D-4.

[52] M. Iglesias Olmedo, T. Zuo, J. B. Jensen, Q. Zhong, X. Xu, S. Popov, and I. Tafur Monroy, "Multiband carrierless amplitude phase modulation for high capacity optical data links," *in Journal of Lightwave Technology*, **32**(4), 798-804 (2014).

[53] D. Qian, N, Cviketic, J. Hu, and T. Wang, "108 Gb/s OFDMA-PON with polarization multiplexing and direct detection," *in Journal of Lightwave Technology*, **28**(4), 484-493 (2010).

[54] K. Kikuchi, and S. Kawakami, "16-ary Stokes-vector modulation enabling DSP-based direct detection at 100 Gbit/s," *in Optical Fiber Communication Conference and Exhibition*, (Optical Society of Ameirca 2014), paper Th3K.6.

[55] M. Chagnon, M. Poulin, S. Lessard, and D. V. Plant, "1 224 Gb/s 10 km transmission of polarization division multiplexed PAM-4 signals using 1.3μm SiP intensity modulator and a direct-detection MIMO-based receiver," *in European Conference and Exhibition on Optical Communication*, (Optical Society of America, 2014), paper PD.4.4.

[56] K. Kikuchi, "Electronic polarization-division demultiplexing based on digital signal processing in intensity-modulation direct-detection optical communication systems," *in Optics Express*, **22**(2), 1971-11980 (2014).

List of Acronyms

ADCs analog to digital converters

ASE amplified spontaneous emission

AWG arrayed waveguide grating

ASIC applications-specific integrated circuit

BNGs broadband network gateways

B2B back-to-back

CAP carrierless amplitude phase

CAPEX capital expenditures

CDNs content distribution networks

CFP C-form factor pluggable

DC datacenter

DMT discrete multitone

DSL digital subscriber line

DSP digital signal processing

DT Deutsche Telekom

EDFA erbium-doped fiber amplifier

FEC forward error correction

FFT fast fourier transform

FFTH fiber to the home

KET key enabling technology

GMPLS generalized multi-protocol label switching

ICT Information, communication, and technology

iFFT inverse fast Fourier transform

IM/DD intensity modulated/direct detection

IM interface module

IP Internet protocol

LC left circular

MILP mixed integer linear programming

MIMO multiple-input-multiple-output

MLR mixed line rate

NE network equipment

NRZ-OOK non-return to zero on-off keying

OEO optical-electrical-optical

ODN optical distribution network

OLT optical line terminal

OPEX operational expenditures

ONU optical network unit

OSNs online social networks

OTN optical transport network

OXC optical cross connect

PCE path computation element

PONs Passive optical networks

QoS Quality of service

RC right circular

RTT round trip time

SD-CON Software defined cognitive optical networks

SRAMs static random access memories

SSMF standard single mode fiber

TCAMs ternary addressable content memories

TDM time division multiplexing

SDN software defined networking

WDM wavelength division multiplexing

XPA XPower Analyzer

4-PAM 4 levels pulse amplitude modulation