SDN/NFV Control Plane Optimizations for Fixed-Mobile Access and Data Center Optical Networking

Bogda-Mihai Andruş

Supervisors:
Professor Idelfonso Tafur Monroy,
Professor Juan José Vegas Olmos and
Dr. Achim Autenrieth

Delivery Date: 10$^{\text{th}}$ October 2017

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

DTU Fotonik
Institut for Fotonik

Abstract

This thesis seeks to leverage SDN and NFV technologies to enable scalable, high performance, resilient and cost effective Fixed-Mobile Convergence (FMC) on three separate segments: access/aggregation, inter but also intra data center (DC) networks for distributed Next Generation Points of Presence (NG-POP - the location in the operator infrastructure centralizing the IP edge for all access types).

From a structural stand point, shared access/aggregation for data transport solutions (i.e., low latency optical cross connects and Wavelength Division Multiplexed-Passive Optical Networks) are proposed, evaluated and assessed for compatibility with the stringent bandwidth, latency and jitter requirements for baseband mobile, fixed and Wi-Fi services. Functional convergence is achieved by integrating universal network functions like Authentication, Data Path Management, content caching inside the NG-POP with the help of SDN and NFV techniques. Comprehensive convergence is demonstrated, for the first time to our knowledge, through the study of a series of use cases highlighting the FMC impact on user experience: seamless authentication and roaming through various network access points (i.e., Wi-Fi, mobile, fixed) as well as improved QoS and network utilization through content caching. To provide survivability for NG-POPs, Network Function live migration with zero downtime is demonstrated as a prevention mechanism for infrastructure failure.

High performance scalable DCs are required by NG-POPs for providing services like content caching, multi-tenancy support, customer applications among others. To develop compatible, cost effective DC solutions, we investigate new high radix DC topologies (i.e., hypercube, fat tree, torus, jellyfish) as well as optimizing control plane operations by adopting SDN. We demonstrate that, for a high performing topology like hypercube, SDN can achieve a 45% throughput increase, as opposed to conventional Spanning Tree Protocol.

Optimizing elastic optical network control plane for DC connections, is another aspect investigated in this work. We propose a modular scalable flexi-grid optical domain controller based on Finite State Machines and a NETCONF/YANG standard northbound interface. The modular structure allows either a centralized or a distributed deployment for on the fly encrypted device management connections. Controller evaluation over networks ranging from 1 to 64 ROADMs show a relatively constant start up and synchronization time in both deployments. These results, together with a modest log scale growth of media channel set up time, validate the scalability of our controller.

Resumé

Denne afhandling sigter mod at udnytte SDN og NFV teknologier til at muliggøre skalerbar, høj ydeevne, robust og omkostningseffektiv Fixed-Mobile Convergence (FMC) på tre separate segmenter: adgang / aggregering, inter og intra data center (DC) netværk til distribuerede Next Generation Points of Presence (NG-POP - placeringen i operatørinfrastrukturen der centraliserer IP-kanten for alle adgangstyper).

Fra et strukturelt standpunkt foreslås, vurderes og valideres fælles adgangs / aggregeringstransportløsninger (dvs. optisk krydsforbindelser med lav latency og bølgelængde-division Multiplexede-Passive Optiske Netværk) som er kompatible med de stringente båndbredde-, latency- og jitterkrav for baseband mobile, fastnet og Wi-Fi-tjenester. Funktionel konvergens opnås ved at integrere universelle netværks funktioner så som autentificering, Data vejs Management og indholds caching inde i NG-POP ved hjælp af SDN og NFV teknikker. Omfattende konvergens demonstreres for første gang, ifølge vores viden, gennem en række bruger tilfælde, der fremhæver FMC's indvirkning på brugeroplevelsen: sømløs godkendelse og roaming gennem forskellige netværksadgangspunkter (dvs. Wi-Fi, mobil, fast) samt forbedret QoS og netværksudnyttelse gennem indholds caching. For at give overlevelsesevne til NG-POP'er, demonstreres en netværksfunktion med nulstilling og uden nedetid, som en forebyggende mekanisme for infrastrukturfejl.

Højtydende skalerbare DC'er er påkrævet af NG-POP'er for at levere tjenester som indholds-caching, multi-forbruger og kundeapplikationer mv. For at udvikle kompatible, omkostningseffektive DC-løsninger undersøger vi nye høj-radix DC-topologier (dvs. hypercube, fat tree, torus, jellyfish) samt optimering af kontrolplanoperationer ved at implementere SDN. Vi demonstrerer at SDN, for en højtydende topologi som f.eks. hypercube, kan opnå en 45% gennemstrømningsforøgelse i modsætning til den konventionelle Spanning Tree Protocol.

Optimering af et elastisk optisk netværkskontrolplan for DC-forbindelser, er et andet aspekt, der undersøges i dette arbejde. Vi foreslår en modulær skalerbar flexi-grid optisk domæne controller baseret på Finite State Machines og en NETCONF / YANG standard for nordgående grænseflader. Den modulære struktur tillader enten en centraliseret eller en distribueret implementering for krypterede enhedsadministrationsforbindelser. Evaluering af kontrolmekanismen, over netværk, der spænder fra 1 til 64 ROADM'er, viser en forholdsvis konstant opstarts- og synkroniseringstid på ca. 920 ms for den centrale og 1150 ms for den distribuerede implementering. Disse resultater, sammen med en beskeden log-skala vækst af mediekanalers opsætningstid, validerer skalerbarhden af vores controller.

Acknowledgements

This thesis is based on the achievement of my research within the department of Photonics Engineering, Technical University of Denmark and with ADVA Optical Networking SE in Martinsried/Munich, Germany.

First of all, I would like to thank my supervisors Prof. Idelfonso Tafur Monroy, Prof. Juan José Vegas Olmos and Dr. Achim Autenrieth for giving me the opportunity to do my Ph.D. in their group, for their support and also their belief in me during the course of my Ph.D.

In addition, I would like to thank my colleague from ADVA: Thomas Szyrkowiec for all the help and motivation to improve my work and expand my knowledge in the field of SDN.

From DTU, I would like to thank, Saul Vasquez for printing this thesis, Lau Suhr for translating this abstract in Danish and all my colleagues from the former Metro-Access group, for their personal support.

To my parents back home, to whom I owe all my personal and professional achievements, I want to thank them for their invaluable faith and moral support. Finally, I want to thank my girlfriend Corina Greșanu for proofreading this thesis but especially for her love, inspiration and patience during all my studies.

Bogdan-Mihai Andruş
November 2017

Summary of Original Work

This thesis is based on the following original publications:

PAPER A

B. Andrus, A. Autenrieth, S. Pachnicke, S. Zou, J.J. V. Olmos and I. T. Monroy, "Performance Evaluation of NETCONF-Based Low Latency Cross-Connect for 5G C-RAN Architectures," Accepted for publication in *International Conference on Transparent Optical Networks (ICTON)*, 2018, (*Invited*).

PAPER B

B. Andrus, R. Martinez, A. Autenrieth, M. Requena, R. Vilalta, B. Le Guyader, X. Grall, S. Gosselin, P. Olaszi, A. Pineda, A. Ladanyi, S. Zou, J. J .V. Olmos and I. T. Monroy, "Demonstration of Next Generation Point of Presence for Fixed-Mobile Convergence," Submitted for publication to *International Journal On Advances in Networks and Services*, issn: 1942-2644, March 2018, (*Invited*).

PAPER C

B. Andrus, A. Autenrieth, S. Pachnicke, J. J. V. Olmos and I. T. Monroy, "Live Migration Downtime Analysis of a VNF Guest for a Proposed Optical FMC Network Architecture," in *Proc. of 17th ITG-Fachtagung Photonische Netze Conference*, Leipzig, Germany, 2016.

PAPER D

B. Andrus, O. M. Poncea, J.J. V. Olmos, and I. T. Monroy "Performance Evaluation of Two Highly Interconnected Data Center Networks," in *Proc. of 17th International Conference on Transparent Optical Networks (ICTON), 2015.*

PAPER E

B. Andrus,V. Mehmeri, J.J. V. Olmos, S. Spolitis, V. Bobrovs, and I. T. Monroy, "SDN Data Center Performance Evaluation of Torus and Hypercube Interconnecting Schemes," in *Proc. of Advances in Wireless and Optical Communications (RTUWO)*, 2015.

PAPER F

B. Andrus, V. Mehmeri, A. Autenrieth, J.J. V. Olmos, and I. T. Monroy, "Evaluation of SDN Enabled Data Center Networks Based on High Radix Topologies," Accepted for publication in *Proc. of International Conference on Systems and Networks Communications (ICSNC)*, 2017.

PAPER G

B. Andrus, A. Autenrieth, T. Szyrkowiec, J.J. V. Olmos, and I. T. Monroy "Evaluation and Experimental Demonstration of SDN-Enabled Flexi-grid Optical Domain Controller based on NETCONF/YANG," Accepted for publication in *Network Operations and Management Symposium (NOMS)*, Taipei, 2018.

Other scientific reports associated with the project:

R1

S. Pachnicke, **B. Andrus**, and A. Autenrieth, "Impact of Fixed-Mobile Convergence, ", in *International Conference on Optical Network Design and Modeling (ONDM)*, 2016.

R2

A. Autenrieth, S. Zimmermann, S. Zou, **B. Andrus** et al., "Report describing results of operator testing, capturing lessons learned and recommendations", COMBO ICT Project Deliverable D6.3, July 2016, found at: `http://www.ict-combo.eu/data/uploads/deliverables/combo_d6.3_wp6_2016-07-22_adva_v1.0.pdf`

List of Acronyms

3R Retime, Reshape, Regenerate

3W WiMax, Wi-Fi, WCDMA)

5G 5th generation wireless systems

AAA Authentication, Authorization and Accounting

ABNO Application-Based Network Operations

ANM Adaptive Network Manager

API Application Programming Interface

AP Access Point

BBU Base Band Unit

BER Bit Rrror Rate

BPSK Binary Phase Shift Keying

BV-WSS Bandwidth Variable Wavelength Selecrive Switch

BV-WXC Bandwidth Variable Wavelength Cross Connect

BVT Bandwidth Variable Transponder

C-RAN Cloud RAN

Capex Capital Expenditure

CCAMP Common Control and Measurement Plane

CC Content Caching / Cache Controller

CDC Colorless Directionless Contentionless

CDN Content Delivery Network

CD Colorless Directionless

CLI Command Line Interface

COMBO COnvergence of fixed and Mobile BrOadband access/aggregation networks

CORD Central Office Rearchitected as a Datacenter

CPRI Common Public Radio Interface

DC Data Center

DWDM Dense Wavelength Division Multiplexing

ECMP Equal Cost MultiPath Protocol

eNB Evolved Node B

EON Elastic Optical Network

EPC Evolved Packet Core

EPS Electrical Packet Switching

EVPL Ethernet Virtual Private Line

FBFly Flettened Butterfly

FEC Forward Error Correction

Flexi-Grid Flexible DWDM Grid

FMC Fixed Mobile Convergence

FPGA Field Programmable Gateway Array

FSAN Full Service Access Network

FSM Finite State Machine

G-PON Gigabit-capable PON

GMPLS Generalized Multi-Protocol Label Switching

I/Q In-phase/Quadrature

ICP Internet Communications Provider

ICT Information and Communications Technology

IETF Internet Engineering Task Force

ILA In Line Amplifier

IMS IP Multimedia Sub-System

IP Internet Protocol

ISP Internet Service Provider

ITU-T International Telecommunication Union - Telecommunication Standardization Sector

JVM Java Virtual Machine

KVM Kernel-based Virtual Machine

LAN Local Area Network

LTE Long-Term Evolution

MBH Mobile Backhaul

MME Mobility Management Entity

MPLS Multiprotocol Label Switching

MSA Multi Source Agreement

NETCONF Network Configuration

NFV Network Function Virtualization

NF Network Function

NG-PoP Next Generation Point of Presence

NH Netowrk hypervisor

NID Network Interface Device

NS3 Network Simlator 3

OAN Optical Access Network

OCS Optical Circuit Switching

ODU Optical Data Channel Unit

OEO Optical-Electrical-Optical

OFDM Orthogonal Frequency-Division Multiplexing

OF OpenFlow

OLS Optical Line System

OLS Optical Line System

OLT Optical Line Terminal

OMS Optical Multiplex Section

ONOS Open Network Operating System

ONU Optical Network Unit

Opex Operational Expenditure

ORoF Optical Ring of Fat tree

ORCHESTRA Optical peRformanCe monitoring enabling dynamic networks using a Holistic cross-layEr, Self-configurable Truly flexible appRoAch

OSPF Open Shortest Path First

OTU Optical Transport Network

P2P Point-to-Point

PCE Path Computation Element

PON Passive Optical Network

PoP Point of Presence

PSK Phase Shift Keying

QAM Quadrature Amplitude Modulation

QoS Quality of Service

QPSK Quadrature Phase Shift Keying

RADIUS Remote Authentication Dial-In User Service

RAN Radio Access Network

RGW Remote Gateway

ROADM Reconfigurable Add Drop Multiplexer

RPC Remote Procedure Call

RRH Remote Radio Head

SBVT Sliceable BVT

SDN Software Defined Networking

SDON Software Defined Optical Networking

SIM Subscriber Identity Module

SIP Session Initiation Protocol

SSL Secure Sockets Layer

STP Spanning Tree Protocol

TED Traffic Engineering Database

ToR Top of the Rack

TWDM Time and Wavelength Division Multiplexing

UAG Universal Access Gateway

uAUT universal AUThentication

uDPM universal Data Path Management

UE User Equipment

UMA Unlicenced Mobile Access

URL Uniform Resource Locator

VLAN Virtual Local Area Network

VNF Virtual Network Function

VoIP Voice over IP

WCDMA Wideband Code Division Multiple Access

WDM Wavelength Division Multiplexing

Wi-Fi Wireless Fidelity

WiMax Worldwide Interoperability for Microwave Access

WR Wavelength Routed

WSON Wavelength Switched Optical Networking

WSS Wavelength Selecrive Switch

WS Wavelength Selected

XC Cross Connect

XML Extensible Markup Language

YANG Yet Another Next Generation

Contents

Chapter 1

Introduction

1.1 Overview of Research Directions

In the area of optical communications, extensive research and development efforts are focused towards extending the flexibility and efficiency of optical networks. The main driver is to better adapt to new emerging bandwidth-demanding services while catering to the needs of end-users as well as reducing the overall cost and complexity of the network. The nature of optical networks has experienced a considerable transformation in the last decade mainly because of the ever-increasing traffic resulted from the rapid adoption of broadband connectivity and the emergence of new bandwidth-demanding and Quality-of-Service (QoS) critical applications and services.

Photonic technologies represent the most suitable solution to support and address such massive bandwidth demands. One capacity improvement technique in optical networks is making use of wavelength division multiplexing (WDM) technology to add ad-hoc new wavelength channels employing fixed modulation formats and bitrate serial interfaces. Nonetheless, it is becoming apparent that this development model may not be able to provide support for the future needs - estimations suggest that global IP traffic will increase nearly threefold over the next 5 years with a compound annual growth rate of 24% from 2016 to 2021 [1].

Bridging the gap between the current infrastructure and traffic demands can only be achieved by a combination of increased spectral efficiencies and improved utilization of resources. Future optical networks are hence envisioned as flexible or elastic, in order to accommodate more efficiently the shortage of bandwidth. Centralized proprietary applications for network control have been around for years. Using a combination of various technologies they mostly require closed plugins to interoperate with network devices.

SDN started off as a more flexible alternative to network control and management by defining open interfaces between a centralized controller and the network devices. Currently, SDN-based system architectures are investigated for deployment in various network segments ranging from data centres to campus, access, metro even core. One of the end goals envisioned for SDN, is to provide end-to-end orchestration for telecommunication services through the use of multiple heterogeneous transport technologies making service-related functions independent from underlying transport-related technologies and administrative domains.

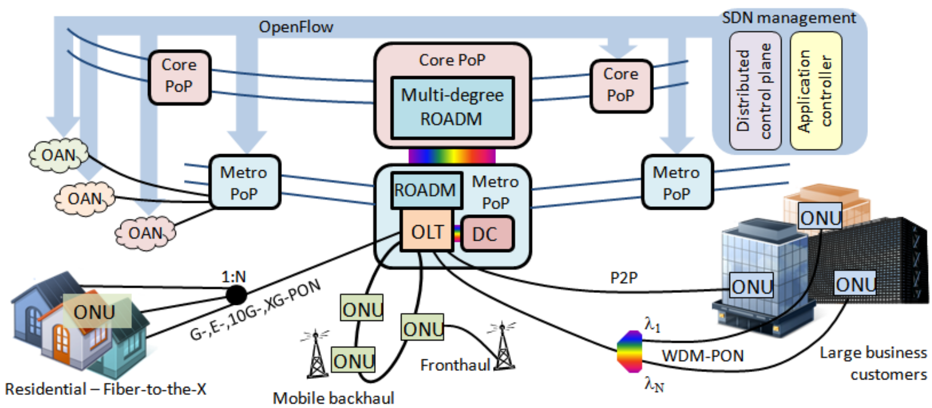


Figure 1.1: Proposed unified control plane of general optical network

Figure 1.1 shows the overview of a proposed heterogeneous data plane and the unified control plane in an optical network separated into access, aggregation/metro and core segments. Next generation points of presence (NG-POPs) facilitate the interconnection between each segment. We identify as one of the main challenges (opportunity) of the aggregation level NG-POP the need to provide support for the heterogeneous access types by acting as a fixed-mobile-wireless service gateway. Therefore, in its structure, components like Optical Line Terminations (OLT) for access networks and Data Centre (DC) features required for hosting mobile Base Band Units (BBUs) and other specific network functions (NFs), are to be included. The core PoP must in turn support and transport large and irregular traffic loads. With the use of SDN-enabled flexible optical transponders (able to operate at different bitrates, modulation formats and wavelength channel spacings), that offer high spectral efficiency on the fly, the variable traffic patterns can be accommodated.

Developments in the area of elastic networks and flexi-grid have the potential to significantly unlock more bandwidth in the core optical network, and furthermore, offer the necessary means for cost effective data center interconnection transport solutions. The FP7 EU Marie Curie project ABACUS funding this work, takes a holistic approach to the development of control and management systems for optical networking, understanding that changes to a specific network segment affects design and operations in other sections. Therefore, the following relevant subjects which are investigated in this work can be divided into 3 main areas: next generation of Converged Fixed-Mobile access/ aggregation networks, intra-Data Center networks seen as an integral part of distributed NG-POPs, and SDN-based control plane optimizations for EONs based on flexible DWDM grid technologies as an interconnection solution for distributed NG-PoPs.

1.2 Next Generation of Converged Fixed-Mobile Access/Aggregation Network Architectures

Access networks have been experiencing an increase in traffic over the past couple of decades without any decline prediction for the currently increasing trend [2]. Convergence of Fixed and Mobile Services is a subject that has undergone a series of research initiatives, however not all ending as successfully as initially expected [3], [4]. A previous approach focused on a baseline IP Multimedia Sub-System (IMS) [4] architecture which plays the role of a common IP interface so that signaling, traffic, and application development are greatly simplified. The developed platform was offering the operation of network agnostic services across mobile and fixed user devices. However, full network convergence was not resolved, which lead to the perception of two individual and parallel IMS deployments.

Further analysis of the practical applicability of the concept within the industry has revealed a lack of demand for the provided services. As a result, the Fixed Mobile Convergence Alliance formed of 20 global telecom operators was disbanded in 2010. The recent ICT COMBO [5] project approach to FMC seeks to redesign the next generation of access/ aggregation networks by proposing a common network infrastructure which can support fixed and mobile network users. In this regard, FMC is developed from a twofold perspective: a structural convergence (a common transport infrastructure) and a functional one (universal functions serving all users). Furthermore, simplified network management and maintenance can be achieved by adopting SDN and NFV paradigms as presented further on.

Another concept envisioned as part of the next generation of FMC architecture is Cloud Radio Access Network (C-RAN) which promises low bit-cost, high spectral and energy efficient mobile broadband Internet access to wireless customers. C-RAN's basic characteristic requires the decoupling of the tight and close connections between the Remote Radio Heads (RRHs) and the Base Band Units (BBUs) which are then centralized in a common and virtualized pool. One of the major disadvantages of the C-RAN centralized model is the high bandwidth requirement between the BBU and the edge RRH needed to carry the baseband I/Q signal [6]. When it comes to real time services like voice or video streaming, network latency alone can have a crucial impact on the user experience. Various radio interface standards define one-way latency constraints for the fronthaul transport with upper limits ranging from 100 us [7] and often placed at 250 us [8] mostly related to maximum supported transmission distance over fiber. Due to strict synchronization requirements (e.g. optimal and secure handover between baseband stations, minimal disturbance on the air interface, frequency stability) a deviation limit (i.e., jitter) is also imposed on the fronthaul, equivalent to 16.276 ns for round trip transmissions. Therefore, the FMC architecture must be built upon high speed, low latency, flexible optical transport for the radio technologies distributed across the edge of the network.

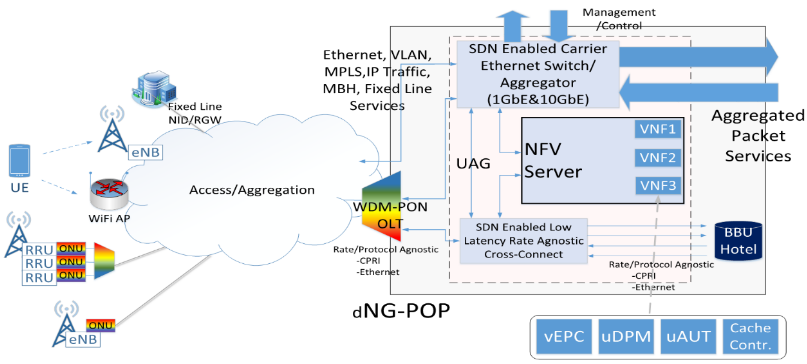


Figure 1.2: Fixed Mobile Convergence system architecture with a shared Access/ Aggregation and an NG-POP. (CPRI: Common Public Radio Interface; UAG: Universal Access Gateway; vEPC: Virtual Evolved Packet Core; uAUT: Universal Authentication; uDPM: Universal Data Path Management)

An overview of the FMC access/ aggregation network demonstrated in this thesis is presented in Figure 1.2. The unified transport network covering the access and aggregation segments needs to be based on technologies able to fulfill the previously identified major requirements related to bandwidth (e.g. higher bit rates above 10 Gb/s) and strict jitter and latency limits. A suitable candidate for a shared fronthaul network, capable of fulfilling the previously stated requirements, is a Wavelength-Division Multiplexed Passive Optical Network (WDM-PON), currently under standardization by ITU-T and FSAN, Study group 15, (ITU-T G.989.2) [9–11].

At the core of the architecture and also in the focus of our research efforts is the Next Generation Point of Presence (NG-POP) which combines structural and functional components and features. The main functional role of the NG-POP is to provide a Universal Access Gateway (UAG) for all connecting users. We identify several building blocks contained in the NG-POP and highlight them in Fig 1.2. An SDN enabled low latency optical cross connect enables the dynamic re-assignment of BBUs hosted in the locally centralized BBU pool to different RRHs at the edge of the network according to various use cases resulting and variable traffic patterns (e.g. tidal effects). The role of the Carrier Ethernet switch is to aggregate all the user connections from the various access points (e.g. fixed, mobile, wireless) onto higher line rate links towards the core network. Moreover, it can also identify and label user connection types and steer them internally for processing by the UAG.

The NFV represents an open and expendable platform which is the key element in the functional convergence. It can host a wide range of Virtual Network Functions (VNFs) like authentication, path management, evolved packet core, caching etc.

Mobile traditional architectures place the Evolved Packet Core in a centralized layout. As the density of user devices covered and requirements of delivered services increases [12], a shift of EPC features towards a more distributed approach, closer to the end user, becomes logical. Improvements on resiliency, scalability and especially service continuity in case of disaster events also act as promoters for the proposed distributed approach. Deploying the EPC in a virtual environment inside the NG-POP has implications on the BBU uplinks and backhaul network traffic which will consequently terminate inside the NG-POP. By eliminating the need for a remote and secure channel between the mobile core and the BBUs (both are locally confined) the ability to develop local in-path network intelligence, monitoring and content caching becomes available. As a result, the NG-POP represents the most suitable platform to incorporate the necessary additional functional convergence blocks.

Resource access control is one of the most important functions in a network, regardless of the access technology. Providing service for all access types (e.g. fixed, mobile, Wi-Fi) and allowing mobility between these connection points results in a need for universal Authentication, Authorization and Accounting (AAA) control in the UAG. The proposed uAUT system [13] is regarded as a base layer functionality which leads to further degrees of functional convergence. Its most relevant roles focus on provisioning policies at the initial phase of network attachment as well as accounting of the service delivery for billing and auditing purposes. The uAUT is able to establish user services, coordinate address assignment and activity accounting for roaming users by accessing and retrieving user profiles stored in a universal subscriber database.

Expected exponential increase in mobile broadband traffic in 5G networks can be addressed with offloading and handover techniques through mobility management (MME). An important objective for our proposed FMC architecture is accomplished by introducing a universal Data Path Management (uDPM) function [14] [15]. uDPM is directly responsible for allowing users to roam between all access types and for redirecting traffic across several types of interfaces. In order to achieve this, uDPM provides a converged subscriber and session management, an advanced interface selection and other route control mechanisms.

Improving QoS for connecting users and reducing redundant network traffic is another scope of the FMC architecture. Locally caching content is proven to have benefits directly proportional with the number of users on the segment. In live deployments of content caching for fixed access networks, it has been shown that traffic can be reduced with more than 30% [16], [17]. Caching in mobile networks however, has not proved to bring any significant improvements [18]. To this end, a unified content distribution system is introduced, creating the basis for a Content Delivery Network [19]. Applying caching functionality in a FMC network can have a multiplicative impact on the reduction of traffic. The system is made up of a Cache Controller incorporating the caching intelligence and a Caching Node storing the cached content as seen in Fig. 2. The functionality is directly linked to the uAUT service for authenticating at a network and service level. uDPM is also responsible for supplying resource information to the Cache Controller regarding UE location (e.g. client ID or IP, cache address, content URL etc) as well as network performance.

1.3 Data Center Networks

While the main trend regarding data centers is focusing on increasing their size and performance, an alternative approach turns towards a geographical distribution of data centers in key places in the network (e.g. NG-POPs), closer to the customers. Installing business critical applications and IT infrastructure closer to the office location is preferred, in many situations, over the choice of a distant central location. Important motivations for geo-distributed data centers lie in reducing latency and cost while increasing efficiency and reliability. Empirical studies emphasize the direct correlation between speed/latency and business revenue [20]. As an example, an additional 100ms delay on the Amazon platform caused a 1% decrease in sales while a 500ms latency in displaying Google search results lead to lower revenues by 20%. This serves as the basis of developing a data center solution that is physically closer to the customers and businesses and acts as a network extension of a centralized data center that can offer similar services as traditional ones. In addition, adopting geo-diversity for data center placement serves as a redundancy mechanism improving system availability in the case of an outage that can impact an entire site.

One fundamental tenant of the envisioned plan for 5G [21] supports the idea of hosting 5G network functions, content and applications in data centers distributed closer to the edge. In other words, operators can develop physical assets closer to the edge network (mobile access/ aggregation) and transform them into distributed data centers. Network slicing is expected to address the needs of dedicated services and particular user groups and provide custom resources in real-time 5G network conditions. Enabling such features requires data center oriented infrastructure. Furthermore, latency is a key requirement for 5G services with current discussions supporting delays as low as $1ms$ [22]. To enable this technological requirement and support the deployment of future expected services like mobile payments, home automation, vehicle connections etc. data centers are expected to be deployed in a more distributed pattern closer to the cell towers.

Deployment of distributed data centers can provide added value not just for 5G services. C-RAN architecture also seeks to apply data center technologies to allow for increased bandwidth, highly reliable, low latency interconnections in BBU pools. Proper placement of the distributed data centers is a crucial step which, if not optimized, can lead to increased costs of providing services instead of decreasing them. According to the proposed FMC architecture presented in the previous sections, we strongly believe that the modular and flexible NG-POP structure represents the most suitable location to incorporate a data center, thanks to multiple reasons. The location of the NG-POP is close enough to the edge network to support the BBU-RRH separation and centralize them in an internal pool. Data center features are necessary not only for the BBU pools but also for the variety of FMC specific network functions which require hosting (e.g. uAUT, content caching, uDPM etc.). Additional client oriented applications like office productivity applications, file sharing, email exchange represent obvious candidates for a distributed implementation. Moreover, improvements in system software would enable geographically distributed data centers to support a wider class of applications with a shift from hosting embarrassingly-distributed applications towards more internal traffic intensive applications (e.g. grid computing, search indexes, cloud gaming, social networking etc.) [23]

Recent projections released by Cisco Global Cloud Index 2016 [24] expect global data center traffic to reach 15.3 ZB by the end of 2020 with a growth rate of up to 4.7 ZB every year. The biggest contributer to this growth remains the east-west traffic patterns inside the data centers which accounts for more than 75% of the total. Such a projection emphasizes the importance of performance for internal data center networks that need to handle the growing amount of packets traveling between servers. The interconnection topology is an important factor in the performance of data center applications which share more and more similar features with efficient on-chip network functions [25]. As a result, highly interconnected topologies initially adopted for on-chip networks in parallel computing have gained the interest for applications inside data center networks [25]. A few relevant topologies which serve as good topological candidates like fat tree, jellyfish, torus and hypercube represent the focus of our work and are represented in Figure 1.3.

Without a proper assessment of the design and an efficient control and management system, data center networks can lead to an increased cost of providing service. When selecting a network topology for designing a data center, important considerations have to be taken into account. A balance between accepted throughput per input (bandwidth), latency, reliability, cost, complexity and scalability has to be achieved. A general indication of network bandwidth, link density and resiliency is dependent on the mathematical analysis of the bisection bandwidth, a cut through the middle of the network. Traversing nodes and links incurs latency. Therefore, the actual diameter of the layout is a good indicator of the resulting latency.

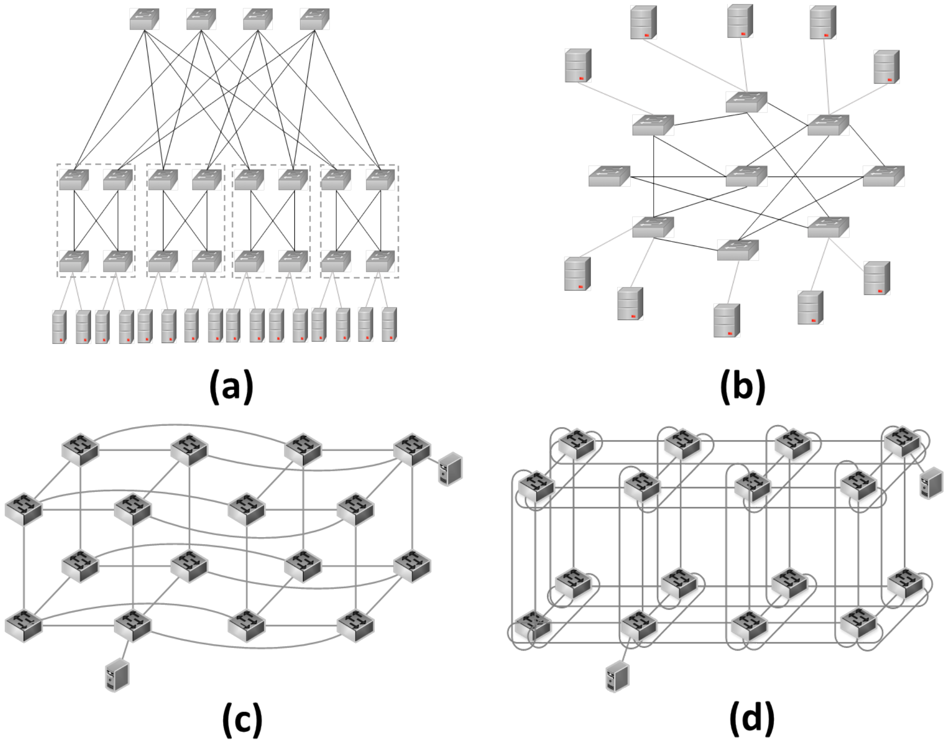


Figure 1.3: Highly interconnected data center network topologies: Fat Tree (a), Jelly-Fish (b), Torus (c) and Hypercube (d).

Reliability of the network is given by the number of redundant paths in the overall layout which contribute to increasing the network's ability to experience local failures without major impact on operations. Scalability, on the other hand, shows how the performance of a system behaves under the action of expanding the network to include new computational nodes. When the number of Top-of-the-Rack (ToR) switches is known, the number of redundant links and interconnection ports impact directly the cost incurred by implementing a certain topology.

A considerable share of the total Capex (capital expenditure) involved with developing a data center is invested in the network. Dedicated switches, routers and load balancers concentrate the majority of the costs related to networking [26]. One of the main advantages brought by adopting an SDN approach is the ability to replace these proprietary dedicated network devices with simplified vendor neutral and cost effective devices. In other words, costs are lowered by substituting expensive nodes which perform complex path computation algorithms and network functions with bare-metal switches that provide fewer features but enable low-cost and flexible alternatives. Infrastructure scalability is another benefit brought by an SDN architecture, with new devices being added more easily to the network. The software controller is capable of scaling to as many network devices as there are in the data center with the use of automation and scripting techniques. Therefore, adding hardware not only simplifies the scaling operations but also lowers the configuration time required. By providing a custom and application-oriented way of managing throughput and connectivity, SDN can improve parallel processing of large data sets (e.g., big data). Policy and security management becomes more efficient with the advent of SDN control which allows for a more efficient way of configuring firewalls and security devices with custom and network-aware security policies. Combining the highly interconnected network topologies with a centralized SDN control plane results in a flexible and intelligent network design capable of taking advantage of the multitude of redundant links comprising such topologies.

1.4 Control and Management of Elastic Optical Networks (EONs) for Inter Data Center Connectivity

One of the key tenants of fixed-mobile convergence focuses on the deployment of a NG-PoP which hosts a distributed data center alternative for new mobile-fixed-wifi oriented services and applications. Evolving towards such an architecture leads to a search for new traffic-demanding solutions for inter data center connections. One expected characteristic of the inter connecting traffic patterns is defined by heterogeneous fluctuations as driven by replication and resiliency requirements of hosted applications and services. As certain applications and services evolve towards a more distributed approach , data center inter connections need to rely on strict transport latency thresholds. Conventional transport networks were designed to cope with traffic growth and provide more complex services than just point to point connections These architecture focuses mainly on over-provisioning the required capacity to guarantee QoS and traffic demands. Such trends have been regarded as the main driver for research carried out in the field of EONs with the goal of providing efficient connectivity solutions. Flexible DWDM grid networks have the potential to increase efficiency and flexibility for optical communications.

1.4.1 Flexible DWDM Grid Optical Networks

Conventional ITU Wavelength Division Multiplexed (WDM) networks [27] are constrained by a fixed 50 GHz spacing grid between the optical signal central frequencies. In order to cope with optical switching and filtering requirements, guard bands are imposed between channels which account for 25% of the fiber capacity. This model has served operators with little degrees of freedom for managing traffic growth requirements which can be accomplished by either increasing the transponder capacity per channel or by moving to a denser grid spacing. As the traffic demands continue to grow and new technologies for improving channel capacity require a considerably increased cost for development, the C-band (the spectrum region between 1.3 and 1.6 1.55 μm with lower losses in optical transmission) is becoming more and more a valuable and scarce resource.

Predicting the evolution of spectrum utilization and bandwidth demands, new overall spectrum efficiency techniques are being investigated. Innovations in optical transmission which combine coherent detection and pulse shaping processing have led to the ability of transmitting higher bit rates into narrower channels (e.g. 100Gb/s over 33 GHz channel) [28]. Further improvements for transmissions above 100Gb/s, especially for longer distance use cases, require a shift to multi carrier signals. Higher speeds of 400 Gb/s or 1 Tb/s occupy multiple slots as the optical waves pass through a fixed filter. As a consequence, not only is a larger channel spacing required, but also a finer granularity when allocating the spectrum. The solution to this problem is the replacement of the 50 GHz fixed grid with a flexible optical grid. By splitting the spectrum and assigning variable segments in increments of 12.5 GHz, a larger variety of requirements can be fulfilled, as seen in Fig. 1.4. In addition, closely packing the channels together and eliminating the inefficient guard bands leads to a much greater spectral efficiency [29].

According to the modulation scheme used, either single carrier (quadrature amplitude modulation - 8-QAM , 16-QAM, quadrature phase-shift keying - QPSK etc.) or multi carrier (O-OFDM -optical orthogonal frequency division multiplexing), flexi-grid super channels occupy variable spectrum sizes. A key component in an EON is a flexible transponder which has the capability of adjusting the optical bandwidth and transmission reach. To achieve this, the transponder generates the transmission signal in variable

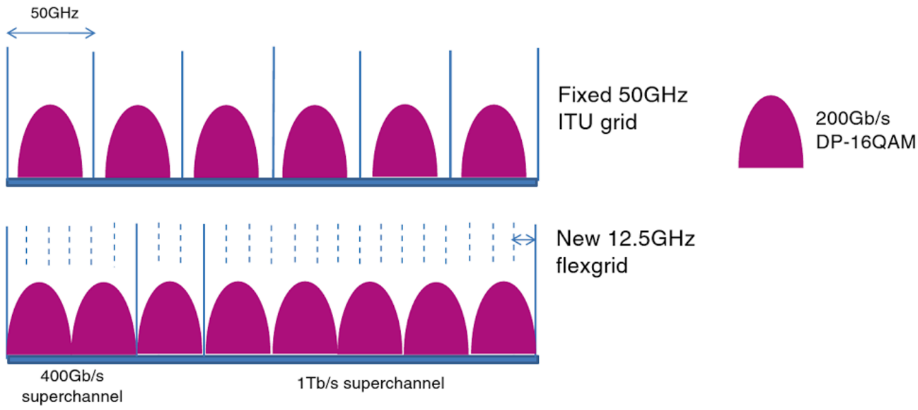


Figure 1.4: 12.5 GHz resolution for flexi-grid allows closer packing channels. Concatenation allows creation of super-channels (e.g. 400 Gb/s, 1 Tb/s).

modulation formats, bit rates, with configurable forward error correction (FEC) and has the ability to shape the optical spectrum. Such a bandwidth variable transponder (BVT) offers the ability to compromise between higher data rates (increased spectral efficiency) on the one hand, and achieving longer transmission distances, on the other hand.

2Following a basic approach, the rule that governs the selection of a modulation format is aimed at minimizing the number of required regenerators while, at the same time, ensuring the highest spectral efficiency. For instance, while selecting 32/64-QAM modulation formats, which provide higher bit rates would be more suitable in shorter range scenarios, schemes like BPSK or QPSK offer higher reaches with the downside of lower bit rates. Nonetheless, the trade-off between reach and bit rate is not linear, side-by-side transponders with double the transmission rate incur a lower cost than installing regenerators for extending the reach of a higher modulated signal.

While a BVT is confined by the ability to allocate a single optical flow and therefore serve only an individual traffic demand, a slice bandwidth variable transponder (SBVT) is capable of generating multiple carriers over which the optical traffic load can be distributed. More specifically, the SBVT can be regarded as a stack of virtual BVTs in a logical association that allocate the data streams over multiple sub-carriers and use multiple modulation formats that differ in spectral efficiency simultaneously (e.g. single or multi carrier).

As the other main component in an EON, the bandwidth variable wavelength cross connect (BV-WXC) is responsible for switching any number of arbitrary sized channels between various input and output device ports. The network of BV-WXC represents the infrastructure over which end-to-end paths are established between BVTs allocating spectral resources to accommodate the generated signals from the transponders side. In a flexi-grid capable network larger channels (e.g. 400Gb/s or 1Tb/s) are allocated the appropriate spectrum width and pass through the BV-WXC filters unmodified. Moreover, lower capacity channels are accommodated by more appropriate adaptable frequency slots with a 12.5 GHz granularity, leaving room for additional signals to be transmitted.

1.4.2 SDN control and management for Inter-Data Center Flexi-grid Optical Networks

Orchestration of inter and intra data center services can be achieved by developing SDN based control and management operations for carrier networks. Abstracting the optical transmission layer and combining it with upper connection layers can lead to a unified control system for various network segments, e.g. access, aggregation, core, data centers. One goal is to provision dedicated resources between distributed data centers with guaranteed QoS. On top of the fundamental data center challenges like energy consumption, location, rack space, administrators also face issues related to variable traffic patterns not just inside data centers but also between them. Some scenarios concerning such dynamic workloads include migration of virtual machines, storage, remote backup, data base synchronizations between geo-distributed data centers. Through the use of a centralized network resource-aware controller, such operations can request temporary connections and release them upon completion, which creates new possibilities for reduced energy consumption.

Conventional Network Management Systems are unsuited to handle flexi-grid technologies to the fullest. On the other hand, SDN architectural principles offer a variety of possibilities when looking to plan, control, and manage flexible network resources both centrally and dynamically. NETCONF [30], an XML-encoded message exchange protocol, is an SDN solution that offers customizable control and management capabilities. Defining and structuring the configuration and state data sent over NETCONF is accomplished with the use of YANG, a data modeling language. In this regard, a flexi-grid network domain controller with a northbound NETCONF interface, together with the YANG model definitions, can provide a standardized control and management solution through the abstraction

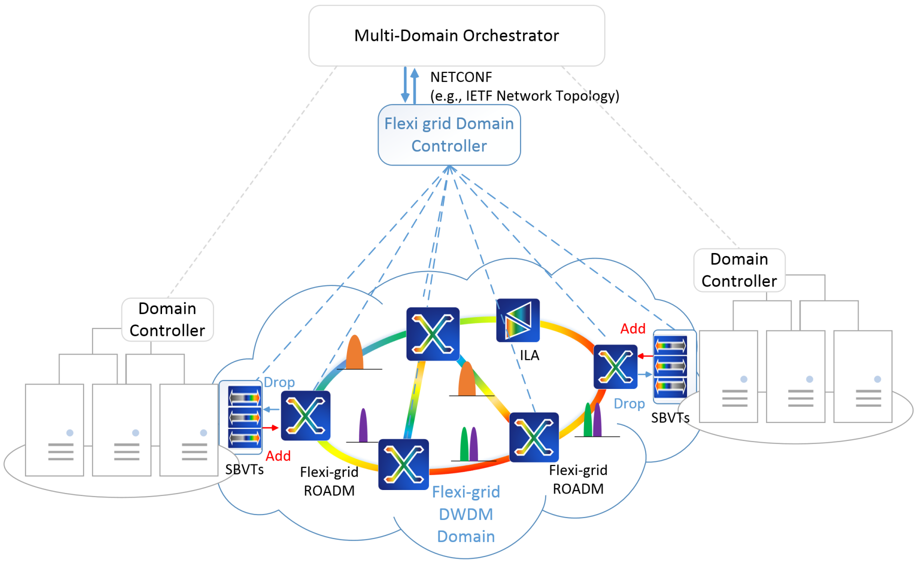


Figure 1.5: Flexi-grid Domain Controller with standard northbound interface (e.g., IETF Network Topology) based on NETCONF protocol.

of the underlining homogeneous network device interfaces. A scenario envisioning a proposed flexi-grid inter-data center domain controller with a standardized NETCONF northbound interface compliant with IETF Network Topology definitions [31] is presented in Fig. 1.5.

One of the SDN design principles uses YANG data models to provide semantic-rich descriptions of the data structure representations of network resources and services. The YANG models can create an imposed agreement not only the between SDN control system and the network elements but also between the system and the northbound application depending on the network architecture (e.g. aggregate or disaggregate).

Both approaches come with advantages and disadvantages and can be exemplified on a DWDM transport network as displayed in Fig. 1.6. On the one hand, an aggregated model, where the entire network acts as a single managed system, is optimized for network operation and can yield higher performance using vendor specific interfaces. End-to-end service orchestration is simplified reducing overall management processes. However, such an approach can affect innovation due to differences in the life cycle of various network elements and also because of strong component interdependency.

On the other hand, in a disaggregated [32] network architecture each

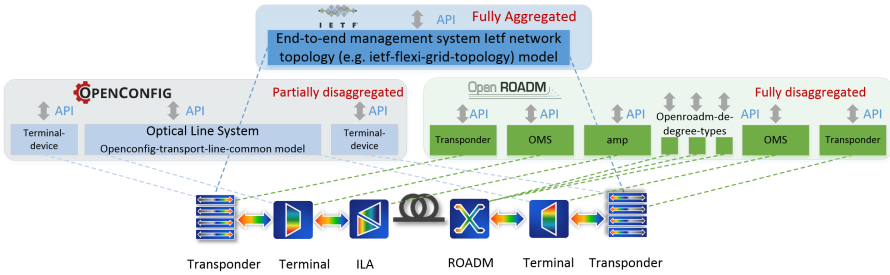


Figure 1.6: Comparison of aggregated and disaggregated network models

component is viewed and modeled as an individual network element. In this context, disaggregation does not refer to decoupling the control plane from the data plane (main SDN tenant). Advantages that arise from such an approach focus mainly on cost optimization and innovation increase. However, operating a disaggregated system not only complicates the central control architecture but also increases the overall synchronization challenge. The interaction between network elements and controller is expected to become more intensive and time-sensitive. As a result, end-to-end service orchestration reaches new complexity levels.

Even though many vendors have embraced the paradigm shift from traditional CLI to programmatic interfaces, the configuration problem converts from syntax to descriptive language models. This means that the models created vary from vendor to vendor and the network operator is left with the same operational problem as before: different processes for different devices that should perform an identical task.

This lack of standard interfaces for network automation has led various work groups and standards organizations to work together to define common configuration data models able to support multivendor network management. Recent initiatives such as OpenConfig, OpenRoadm or IETF CCAMP provide steps in the right direction with respect to optical interoperability and unified management.

OpenROADM Multi-Source Agreement (MSA) [33] focuses on opening up traditionally proprietary ROADM systems for SDN development. The solution adopted by OpenROADM follows a fully distributed approach in which every element in the network is modeled individually, in detail, from tunable transponders to ROADM components (e.g. pluggable optics, degrees, ODU/OTN/multiplex interfaces etc) as seen in Fig. 1.6.

Different DWDM service templates in Remote Procedure Call (RPC) format are also defined in detail. Both Colorless-directionless (CD) and colorless-directionless-contentionless (CDC) options are supported with the assumption that the ROADMs are flexibly configured by an SDN controller. On top of the device templates, OpenROADM also provides an example of network abstraction showing how device and network layer can associate. However, OpenROADM leaves the choice of network abstraction up to the network provider. Even though in the current release the specifications are tailored for a 50GHz fixed grid and 96 wavelengths [34], the upcoming version is planned to offer support for flexi-grid [35].

OpenConfig [36], an informal work group of traditional carriers and ICPs, has similar goals of compiling a consistent set of vendor-neutral data models. Their approach differs however, by adopting a partially distributed network model as seen in 1.6.

Due to rapid rate of innovation in coherent DWDM transponders, a solution to decouple the transponder from the rest of the optical line system allows the network operator to take advantage of the best transponder at any given time. First of all, elements comprising an optical transport line system (e.g., In Line Amplifier - ILA, ROADMs) are grouped and defined in a yang module (i.e., transport-line-common.yang). The OLS module defines the supported ROADMs (incl. CDC ROADMs, WSS and Dynamic Gain Equalizer) as configurable switching elements with input and output ports as well as add/drop ports for directing portions of the optical spectrum to/from the appropriate degree.

Secondly, OpenConfig describes an optical terminal device for a DWDM transport that offers support for optical channel configuration such as wavelength/frequency, power, BER and operational mode. The operational mode structure serves only as a placeholder for vendor-supplied information such as symbol rate, modulation or FEC needed by an EON. However, in the current released version, the models lack support for flexible channel spacing assignment.

The IETF CCAMP working group has released a data model that represents, retrieves and manipulates elements from a flexi-grid capable optical network as a whole [37]. As opposed to the previously presented partially or fully distributed approaches to network modeling, IETF focuses on a more aggregated method dealing with an overview of the optical network topology combined with the underlying physical layer. In this regard, the model representation identifies the most relevant WSON and flexi-grid optical components, parameters and their values and divides them in two sub-modules. On the one hand, an optical Traffic Engineering Database (TED) defines the following elements: - flexi-grid capable nodes: abstract ROADM model with input/output ports and internal port connectivity matrix; - transponders: node augmentations with the role of a tunnel termination point having variable FEC and modulation; - sliceable transponders: defined as a list of regular transponders - links: connections with source and destination nodes and ports that define additional flexi-grid attributes like base frequency, maximum nominal central frequency, allowed frequency granularity and slot width granularity.

On the other hand, the model also describes a media channel structure that is used to create paths from source to destination transponders through a number of intermediary nodes and links. Source and destination nodes can also be specified. The media channel represents both the topology (list of intermediary links and nodes referenced in the optical TED) and the resource it uses (frequency slot [38] - central frequency and slot width). With the scope of maintaining consistency in the configuration data, every element in the TED is assigned a reference which is called by the medial channel.

Additional traffic engineering configuration and monitoring capabilities (e.g. delay, bandwidth metrics, link statistics etc.) are provided in the IETF base models imported and augmented by the flexi-grid YANG. Since this presents a more abstract view of the network topology, any other physical layer parameters regarding optical channel interfaces for DWDM applications are tackled and defined in a different project [39].

1.5 Problem Statement

Conventional fixed, mobile and Wi-Fi networks have evolved separate of each other and, as a result, implementation, structure, equipment as well as network functions and their distribution in the network are very diverse. Therefore, not only the cost of developing (Capex) the separate network types is cumulative but also the operational cost and complexity are increased (Opex). As the mobile subscriber demands (e.g., streaming, real-time communication, social media) lay the foundation for mobile broadband, the average bandwidth offered by the current wireless networks is not comparable to the capacity provided by fixed networks. A lack of seamless and transparent UE mobility and authentication, while roaming between Wi-Fi and mobile networks, impairs from efficiently using access resources. As network functions on both mobile and fixed networks are currently restricted to specialized costly devices (i.e., authentication, service gateways, path management etc.), NFV provides a flexible and scalable alternative to achieving the same functionality on commodity hardware at much lower costs.

Moreover, the existing optical network infrastructure considered as an essential element in the structural FMC, does not have a clear interface to the applications and services spanning across the various domains. Network operators typically lack a mechanism to adjust optical configuration parameters based on the variable conditions of the network or applications and services running on top of them. In this regard, it is difficult to configure these parameters according to predefined policies and to reconfigure them to respond to faults and load changes. In addition, such configuration techniques are done on a segment-by-segment or even device-by-device basis constricting the ability to apply network wide services and applications. Cross layer optimization can be achieved by incorporating the reconfigurable optical device features into the SDN sphere, offering a solution to design centralized control systems for both optical and networking layers.

The NG-POP introduced in this work is foreseen as the location in the operator infrastructure centralizing the IP edge for all access types (mobile, fixed, Wi-Fi). Our developments focus on using SDN and NFV technologies to enable scalable, resilient and cost effective FMC for access/ aggregation networks and improve performance for inter and intra data center (DC) networks required by the deployment of distributed Next Generation Points of Presence (NG-POPs).

1.6 State of the Art and FMC Related Work

As the concept of FMC is extensive and still evolving, initial research focused on a baseline IP Multimedia Sub-System (IMS) [4] [40–42], some also demonstrating a practical deployment [43]. Originally [42], a different set of technological enablers were envisioned for FMC architecture: Session Initiation Protocol (SIP), IMS, VoIP, UMTS, Unlicensed Mobile Access (UMA). Even though some telecom providers previously offered FMC based services using proprietary methods, other research groups like [44] started deploying the first standard oriented 3W services (WiMax, Wi-Fi, WCDMA) on FMC. Research on FMC spans from developing new integrated management architectures [40] to individual hardware terminals using a combination of WiMax and Wi-Fi standards [45].

Other scientific reports associated with the project, and part of the

COMBO consortium, like **R1** , offer a novel new approach on FMC analyzing the cost and performance implications. In addition, the role of optical networks in FMC is introduced and defined by [46] Using SDN principles to orchestrate the automatic provisioning of Fixed Mobile Services is also demonstrated [47].

As SDN is considered to be one of the key enablers for control and management of FMC, the following state of the art documentation details on the latest related research on SDN and its applicability to various components of FMC.

1.6.1 Recent Developments in SDN

Works like [48–53] provide a good overview of general SDN principles and recent developments. In terms of extending the SDN and network virtualization principles and techniques into the optical domain, surveys like [54, 55] cover not only industrial development efforts and research directions but also academic oriented projects. A more in-depth analysis of the difficulties and benefits of a centralized optical control plane are highlighted in [56].

On the one hand, they present how transport level applications like virtual optical network slicing, benefit from a centralized optical SDN controller by setting up end-to-end optical light paths according to bandwidth, latency and jitter restrictions. On the other hand, the researchers acknowledge that the complexity of the control plane increases as more constraints like signal transmission range, amplification, bandwidth granularity and availability, light path routing/switching and configuration timing have to be accounted for.

Considering such challenges, researchers in [57] propose OpenFlow extensions for underlying optical layer abstraction which define general flows across heterogeneous optical transport technologies. Researchers in [58] developed an SDON controller that is capable of creating virtual optical networks through network slicing. Among the demonstrated capabilities we highlight the ability to facilitate the northbound applications to access topology and network information and configure network devices.

Providing a centralized control plane that incorporates the packet and optical network capabilities alike has proved to facilitate control and management of cross domain services [59, 60]. Such developed systems, which aim to solve unified IP and transport network control for a significantly reduced Capex and Opex, define an abstraction level of the underlying network and, based on OpenFlow self-developed extensions, allow the generalization of flows/paths across transport-IP domains. According to our research directions, SDN-related state of the art documentation will focus on the following sub-domains: converged access/ aggregation optical networks, Data Center networks (topologies) and flexible WDM grid optical networks.

Research on SDN enabled Access/Aggregation Networks

Works like [61, 62] explore the applicability of SDN paradigm to access networks. More precisely, the researchers describe use cases which leverage the SDN principles to provide new network services (Dynamic Service Re-Provisioning, On-Demand Bandwidth Boost etc.) or enhance existing ones (e.g. improved Ethernet virtual Private line - EVP -, Service Protection without proprietary mechanism etc.).

Most relevant research directions in this field, focused more on technical implementations, are based mostly on OpenFlow self-developed extensions for supporting either G-PON or even WDM-PON technologies [63–66]. For example, H. Yang et al. in [63] demonstrate service-aware flow scheduling on an OpenFlow-enabled PON testbed.

Furthermore, K. Kondepu et al. in [64] present the development of an SDN-enabled TWDM-PON aggregation node containing an FPGA based OLT, an OpenFlow switch and an internal SDN controller. The purpose of the node is to provide a common converged architecture between the access and aggregation networks thus preserving an overall service level [67].

Research directions towards RAN capable optical transport solutions enhanced by SDN platforms have also been under investigation. Ahmad Rostami et al. [68] have designed and demonstrated resource orchestration across DWDM optical transport for RANs based on SDN. The authors achieved a multi-domain resource orchestration through hierarchical SDN control architecture across transport and RAN layers instantiating end-to-end services like elastic mobile broadband service.

Furthermore, abstraction models for transport resources are developed and assessed by M. Fiorani et al. [69] with 5G service delivery use cases in mind. The presented models are capable of enabling efficient resource orchestration while maintaining the implementation complexity at a reduced level. The author's main goal is to use the models in a transport SDN controller to provide the optimal level of abstraction of optical transport networks in the area of 5G technology development.

Perfromance Evaluation of Highly interconnected SDN-enabled Data Centers

Research related to highly interconnected Data Center networks can be divided into two areas: development and applicability of new topologies and path calculation algorithms. As network topologies are concerned, hybrid interconnections are proposed and evaluated like Optical Ring of Fat tree (ORoF) [70] which offers a finer-grained scalability solution than the regular Fat tree. Upscaling can be accomplished, according to the authors, by either the number of fat tree modules, the number of pods in each fat tree module or increasing both the number of modules and of the pods at the same time.

G. L. Vassoler et al. [71] exploit graph theory to propose a twin-graph-based topology for improved scalability, resiliency and cost. The authors demonstrate that, while performance is kept at a comparable level to the highly interconnected hypercube, the scaling costs have a lower growth rate than hypercube and other studied topologies.

W. Renqun and P. Li [72] propose a Hyper-DC topology which, through mathematical analysis of latency, switch- and link-complexity, promises to provide an efficient solution compared to other proposed topologies like Benes topology [73], 2-dilated flattened butterfly [74] and 2-dilated hypermesh [75]. On the path computation and routing algorithms for DC networks side, SDN has enabled a more rapid development and simplified evaluation of new algorithms due to centralized path computation modules in the controller.

Cosmin Caba presents in his PhD thesis [76] the analysis of an algorithm for offloading elephant flows from the electrical packet switching (EPS) links to the Optical Circuit Switching (OCS) links in hybrid OCS-EPS Data Center topologies. Using mininet for emulating a couple of test topologies (e.g. ring and flattened butterfly FBFly) the author showed that the developed algorithm combined with the strategy of adding shared optical circuits yields a 21%-57% increase in throughput for the ring and a 8% -28% for the FBFly.

Authors in [71] employed an SDN-enabled virtual emulation environment based on an OpenFlow controller to test different routing algorithms (e.g. OSPF - Open Shortest Path First, ECMP - Equal Cost MultiPath Protocol) on a set of various types of topologies.

SDN-based control and management of EONs

Flexible WDM grid optical networks have been regarded as an interesting research topic in the past few years. Several academics and industrial bodies (system vendors and network operators alike) have focused resources on prototype development and demonstrations of control and management systems as well as APIs for flexible optical devices (e.g., transponders) as well as for EONs [77–87].

M. Dallaglio et al. [77], as part of the work being done for the EU Horizon 2020 ORCHESTRA project [78], present the development of control and management system for SBVT with variable modulation, FEC, Bitrate, channel frequency. The SDN protocol of choice for their work is NETCONF and the configuration and management data is modeled in YANG, while the controller architecture is based on Application-Based Network Operations (ABNO) [88]. The authors demonstrate the system capabilities using a pair of emulated transponders developed in C language and running a version of ConfD NETCONF server.

V. Lopez et al. [82] proposes, in broad strokes, in the context of the EU FP7 IDEALIST project, an architecture for an Adaptive Network Manager (ANM) that can operate EONs which is based on 3 main components: an Active Path Computation Element (PCE) for calculating and setting up light paths, and SDN controller which provides a unified view of the network elements and links and acts as a common OpenFlow interface for the devices and am ABNO controller. Use cases like dynamic bandwidth allocation for variable traffic changes, periodic defragmentation for improved bandwidth allocation or multi-layer restoration lead to their chosen ANM architecture.

L. Liu et al. [81] demonstrate a control plane mechanism for spectrum sliced EONs based on the OpenFlow protocol which has the ability to establish end-to-end paths and offload IP traffic. OpenFlow extensions for EON support are developed to carry information like bit rate, frequency slots, modulation format etc. In addition, the network controller (e.g NOX) functionality is enhanced to perform routing and spectrum assignment. With the use of a testbed composed of multi flow bandwidth variable transponders and optical cross connects, the authors demonstrate that an end-to-end path establishment time spanning over more than 3 hops outperforms a GMPLS-based control plane. The performance tests prove that for larger networks a centralized control plane can offer significant improvements over the increased GMPLS signalling times and therefore provides a better solution to scalable networks.

J. Yin et al. [83] take the control plane of EONs one step further and focus their efforts towards the development of a virtual Software Defined EON (vSD-EON) Network Hypervisor (NH). The NH acts as an intermediary which interprets the OF messages received from the northbound ONOS controller into OF with Optical Transport Protocol Extensions for the BV-WSS' running OF agents. The authors evaluate the hypervisor in a full "full-stack" implementation and show that a data connection carrying a video stream over an EON composed of two Optical Transmission Chassis and several 1x9 BV-WSS' is recovered in 2 seconds from the moment of an optical link failure. The data plane restoration is accomplished by the NH in a transparent manner to the northbound controller.

1.7 Beyond State of the Art

This section summaries how the overall scientific results and technical achievements described in this Ph.D. thesis have significantly contributed to current state of the art. Research presented in this thesis fall into three main categories: next generation of Converged Fixed-Mobile access/ aggregation networks, performance improvements for Data Center networks and centralized control plane for EONs based on flexible DWDM grid technologies.

In this regard, **PAPERS A, B** and **C** focus on the aspect of next generation access/ aggregation network architectures. In the view of EU FP7 ICT COMBO project, a unified access and aggregation network architecture allowing fixed, Wi-Fi and mobile networks to converge is developed and demonstrated in this work.

PAPER A analyzes the requirements of an optical transport solution for C-RAN architectures and proposes an SDN-enabled low-latency optical cross connect allowing dynamic scheduling of mobile network resources. We evaluate its performance metrics like one way transmission latency between , jitter and switching time and show that the optical XC complies with the C-RAN requirements. In addition, we develop a configuration API based on NETCONF Protocol and a custom configuration data model written in YANG. The API enables OpenDaylight SDN controller to incorporate the control plane.

In **PAPER B** we present the development of a fully integrated FMC setup with a focus on a Next Generation Point of Presence (NG-POP) which acts as a common subscriber IP edge for fixed, Wi-Fi and mobile access/aggregation networks. At the center of the setup, the NG-POP has a novel architecture developed by combining SDN and NFV concepts. To this end, key universal network functions like Data Path Management, user Authentication, mobile Evolved Packet Core are virtualized and extended to inter-work in order to provide a seamless user experience. We carry out and describe for the first time, to the best of our knowledge, relevant use cases for a converged architecture on a live test bed composed of a transport access/aggregation network infrastructure (e.g. WR and WS WDM-PON), wireless service networks (e.g. LTE and Wi-Fi) and a unified NG-POP.

Seamless user authentication and data connectivity is demonstrated as test user devices roam between multiple LTE and Wi-Fi networks. A content delivery service is also demonstrated where user experience is improved by predicting user mobility (e.g. exchanging messages with the uDPM element) and caching relevant content (e.g. video streams) closer to the user. In addition, we execute performance tests of the unified demonstration setup to evaluate the end-to-end client connection over the various access methods (e.g .fixed/Wi-Fi/mobile) in terms of bandwidth and latency. Through this work, we demonstrate that our proposed FMC concept architecture which is divided into structural convergence (a common optical transport infrastructure e.g. WDM-PON, optical XC switches) and functional convergence (universal access gateway functions centralized in a NG-POP offering converged control mechanism for both fixed and mobile networks) can provide a valid solution for next generation access/aggregation networks and even serve as a reference point for 5G standardization.

Taking the research further, in **PAPER C** we introduce a new use case by providing an NG-POP failure protection mechanism defined by an automated VNF live migration with zero downtime. Using an SDN controller to provision an optical path between the source and destination NFV servers, for the duration of the migration, we achieve a zero downtime functionality.

PAPERS D, E and **F** cover the topic of Data Center networks performance improvements. On the one hand, in **PAPER D** we focus on an analytical evaluation of cost and complexity of modeling DC networks based on a highly interconnected topologies (i.e., hypercube and torus). The analysis of abstract metrics for the two chosen topologies is supported by performance measurements conducted on virtualized simulation test beds. The aim of the tests is not only to compare the performance of the studied topologies but also to asses the scalability effects. Furthermore, in **PAPER E**, we prove that by employing a decoupled control plane and centralizing the topology knowledge inside an SDN controller, network performance can increase by up to 45% compared to conventional switching algorithms. In

PAPER F, we extend our research in order to conduct the same analytical and practical evaluation for new topologies (i.e., jellyfish and fat tree) and offer a comprehensively documented comparison of the current and previous studied topologies.

PAPER G describes our work related to modeling of configuration data for flexible DWDM grid capable optical networks with the aim of optimizing elastic optical network control plane. We develop a modular scalable flexi-grid optical domain controller based on Finite State Machines and a NETCONF/YANG standard northbound interface [89, 90]. The modular structure allows either a centralized or a distributed deployment for on-the-fly encrypted device management connections. We demonstrate the scalability of both deployments using emulated networks with up to 64 ROADMs and two SBVT, by identifying and measuring two performance metrics: start-up time and media channel configuration time.

Chapter 2

Description of Papers

This thesis is based on a set of articles already published or submitted for publication in peer-reviewed journals and conference proceedings. These articles represent the developments and results obtained from our work in the field of control and management systems for optical networking technologies. The covered topics are grouped in three main categories. **Papers A** to **C** focus on the role of Fixed-Mobile Convergence in next generation access/ aggregation networks. **Papers D** to **F** evaluate topologies and control plane optimizations for high performance data center networks. Work in **Paper G** relates to control plane centralization and standardization for heterogenous device interfaces in the EONs domain.

2.1 SDN/NFV Enabled Fixed-Mobile Convergence in Access/Aggregation Networks

Papers A to **C** investigate aspects related to next generation access/ aggregation network architectures. In the context of this work, pertaining to the ICT FP7 COMBO project, a unified access and aggregation network architecture allowing for a structural and functional convergence of fixed, Wi-Fi and mobile traffic is developed, demonstrated and evaluated here.

From a structural standpoint, **Papers A** and **B** contribute in part, to analyzing the performance requirements of an integrated access and aggregation transport solution for the combined traffic types (i.e., fixed, mobile, wireless). C-RAN adoption, defined by the BBU-RRH decoupling and BBU centralization, is identified in **paper A** as a central component in the fixed mobile structural convergence. In order to bridge the technology gap and allow the integration of mobile baseband traffic with the other traffic sources over the same access infrastructure, versatile and optical based transport solutions are considered.

In this regard, in **Paper A**, we propose and develop the model for an SDN-enabled low-latency optical-electrical-optical (OEO) cross-connect solution. The aim of our prototype is to establish and switch the delay and jitter sensitive connections between BBUs and RRHs and allows dynamic scheduling of access network resources. Using a lab test setup we evaluate the previously identified performance metrics (i.e., bandwidth, transmission latency, jitter) displayed by our OEO cross-connect and compare them with the C-RAN requirements. The measured one way transmission latency between 2 - 5 ns for tested traffic up to 10 Gb/s not only complies with minimum C-RAN related specifications of 100 us but also proves a minimum impact on transmission. Furthermore, jitter components present improved characteristics due to the triple R (i.e., Retime, Reshape, Regenerate) functionality of our OEO. An effective delay required by the switch to reconfigure the internal cross-connect and reestablish the optical link was measured at an average of 252 ms in a lab setup environment. Even though jitter was not affected, loss of data occurs during switchover and ongoing connections would be dropped as a result. However, the benefits of counteracting daily traffic tidal effects and other foreseen traffic pattern variations that lead to improved energy efficiency, outweigh a few brief local service interruptions needed for dynamic fronthaul network reconfiguration and optimization. In order to enable such dynamic optimization scenarios, we develop here a software agent based on NETCONF protocol and a custom YANG model, which allows us to decouple the device control plane and integrate it inside an SDN controller (i.e., OpenDaylight). The SDN enabled model is further used in **Paper B** and **Paper C** to enable functional related convergence scenarios.

In **Paper B** we present our developed holistic and integrated FMC setup as the collaborative effort of multiple vendors and operators to achieve functional and structural convergence. As part of the structural convergence, optical WDM-PON networks (i.e. WR and WS) act as a shared access/ aggregation transport medium on top of which mobile, fixed and wireless connections are established between the UEs (e.g., mobile phones, laptops etc.) and a Next Generation Point Of Presence (NG-PoP). The central point of the setup, the NG-PoP, seeks to leverage the advantages of SDN and NFV concepts to achieve functional convergence. Therefore, the

NG-PoP, which comprises of an NFV server, an SDN enabled carrier Ethernet switch and the optical cross connect, acts as a common subscriber IP edge for fixed, Wi-Fi and mobile access/ aggregation networks. To this end, key network functional modules like Data Path Management, user Authentication, mobile Evolved Packet Core (EPC) are developed and connected to allow function chaining for a seamless user experience. Multiple scenarios are envisioned and demonstrated. The universal subscriber and user authentication system was built on top of a RADIUS server, which actively communicates to the EPC module to share information on user profiles (e.g., credentials, content permissions, billing information etc.). The system, allowed us to showcase the user imperceptible authentication while switching through all access methods and using the SIM card as the only UE identification. The second use case, in close connection to the first one, focuses on the data path management functionality of the NG-PoP which deals with the actual UE connection handover when roaming from one network to anotheras the next natural step after authentication. While streaming content from a test server, the UE switches between several network access types (i.e., between two LTE networks, from LTE to WiFi and vice versa) while the connection remains active un uninterrupted. Intelligent content caching scenarios are also demonstrated, where local cache servers store temporarily highly demanded streaming content and save network bandwidth utilization from the core segment. The work is consolidated by executing performance tests (i.e., connection bandwidth and latency) on user connections provided by the integrated setup. This work demonstrates, to the best of our knowledge, the first fully integrated demonstration of a structurally and functionally converged fixed-mobile access/aggreagation network.

Further use cases, in the context of the previously developed fixed mobile converged setup, are conceived and presented in **Paper C**. Live migration of individual Virtual Network Functions (e.g., virtualized load balancers, content servers, security gateways etc.) between different NG-PoPs can serve as an energy efficiency optimization or as a protection mechanism in case of an early failure detection. In order to demonstrate this scenario, we evaluate an automatic SDN provisioning of an optical path through an OEO cross-connect model (i.e., developed in **Paper A**), between two NFV servers running remote OpenStack Compute hosts with KVM virtualization. Step one of our test application requests a data path between the servers through OpenDyalight which controls the optical switch through a NETCONF interface. Step two, triggers the migration request to Open-

Stack Nova module controlling both Compute Nodes by sending a message containing the source and destination hosts and the ID of the guest machine, the VNF. Zero downtime during the migration was accomplished in this scenario. Moreover, rerouting the connection on a different data path during the migration, a scenario which caused by the simulation of a link failure, resulted in a 22% increase in migration delay and an average of 2.1 s downtime.

2.2 Performance Evaluation of Data Center networks

In **Papers D**, **E** and **E** we focus on the evaluation of data center networks from 2 points of view. Firstly, we investigate the topology influence on cost and network performance. Furthermore, we propose and evaluate control plane optimizations with the scope of lowering costs and, at the same time, increasing performance.

In **Paper D**, we present an initial analysis of two highly interconnected topologies (i.e., torus and hypercube) focusing on abstract metrics like bisection bandwidth, diameter, average distance or node order and their implications on resulting throughput, latency, resiliency, complexity, cost etc. A comparative infrastructure cost analysis for implementing the packaging strategies reveals that there is a shift between the two, with the torus being more costly than the hypercube for networks below 64 nodes and the hypercube having a faster growth rate beyond this limit. This conclusion is based on the total number of links, bisection width and node order. As a next step, we replicate the topologies in a network simulator (i.e., NS3) and subject them to uniformly distributed random traffic patterns routed by shortest path algorithms. The selected injection rate is also identified and configured at 30% of the link capacity, below the saturation limit of the studied topologies. The scope of the simulation is to evaluate the scalability of both networks and also compare them with each other. From our simulations we are able to extract performance metrics like throughput, latency and loss rate and characterise the performance of our tested networks. As expected, the growth of the networks and the increase of average node distance lead to a decrease in connection performance. A decline in throughput per connection of about 5% for the hypercube compared to 16% for the torus was measured when the size of the network was increased by a factor of 32 (i.e., ranging from 16 to 512 switches in the tested networks).

Similar behaviour is demonstrated in latency observations and number of lost packets.

Our research on data center networks is extended in **Paper E**, where we focus on performance improvements by optimizing control plane mechanisms and adopting an SDN approach to traffic routing. With the help of a network emulator (i.e., Mininet) and an SDN controller (i.e., Floodlight) we are able to decouple the control plane from the data plane, using OpenFlow protocol, and evaluate the torus and hypercube as the networks scale from 8 to 256 switches. Performance of the SDN frmaeworks for both topologies is also compared to conventional Spanning Tree Protocol (STP). Similar parameters investigated like throughput, latency, jitter and number of lost packets and from the results we conclude that the throughput is higher by roughly 45% for the SDN test cases with 256 nodes compared to STP. An improvement of 13ms in packet delay is also observed in the SDN test cases. Loss rate is also reduced considerably for both topologies with at least 7% by SDN. Connection stability is improved with SDN adoption by a reduction of jitter with 50ms, measurement that applies to networks of 256 switches. This work proves that SDN architecture can bring a significant performance boost in both torus and hypercube, by taking advantage of the multitude of redundant links. The comparison between the topologies remains similar and applies to both studied frameworks. In this regard, hypercube presents superior performance for larger networks, above 64 nodes.

Paper F builds upon our previous work and offers an overview of the total results and lessons learned. We extend our previously presented mathematical analysis of high radix topologies (e.g., Torus, Hypercube) from **Paper D** with regard to indications on performance, cost, latency and complexity of implementation to include new topologies undergoing intense studies in current data center networking research (i.e., fat tree and jellyfish). Using a similar testing environment (i.e., NS3) and configuration parameters (i.e., 30% injection rate, uniformly distributed random traffic and shortest path routing protocol) we demonstrated that the Fat Tree was the lowest performing topology tested with a decrease of 70% in throughput when scaling the networks from 8 to 512 nodes. Additional SDN versus STP comparison is carried out for the jellyfish topology. The results show a two-fold increase in performance for the SDN setup with 120 switches, with the difference slowly decreasing as the network scales up to 256 nodes. The average packet delay was considerably lower in the SDN scenario, with a small dependence on the number of switches employed and presenting less than 1/6th of the delay measured with STP for networks with more than

150 switches. Network jitter and loss rate were also more favorable in the SDN deployment with a reduction in jitter varying from 7% to 33%, Packet loss between the two systems remain within a 2% difference range independently of the number of switches. Our emulation results demonstrate that the SDN deployments based on the studied topologies torus, hypercube and jellyfish, considerably outperform the STP implementation in throughput, latency, jitter and loss rate. This work strengthen the arguments referring to how SDN can be used to leverage the multiple redundant paths in a network to achieve increased performance and more efficinet network utilization.

2.3 Flexi-grid Optical Network Domain Controller based on NETCONF/YANG

Paper G describes our work related to modeling of control plane configuration data for flexible DWDM grid capable optical network devices (i.e., ROADMs, transponders, sliceable transponders, links and media channels).

We develop a scalable NETCONF/YANG flexi-grid optical domain controller with a modular architecture and evaluate its functionality over a network testbed composed of two physical Slice Bandwidth Variable Transponders and an emulated flexible DWDM grid network. We develop the software following a modular architecture based on Finite State Machines (FSMs) which allows us to define an asynchronous, non-blocking communication between the modules. The FSM structure of the Network Driver allows for message pipelining and the resulted multi-threaded and fault tolerant implementation presents good scaling capabilities. The modular aspect allows the flexibility to deploy the controller in either a centralized or in a distributed state. In the centralized deployment, all modules (i.e., Network, Device, Southbound and Northbound Interfaces) are deployed in a single application (i.e., JVM) targeted for a ISP central office. The distributed option decouples the device agents which are spawned in the same location as the physical devices, (i.e., Point of Presence - PoP - hosting the network ROADMs). One major advantage from this approach would be on-the-fly SSL encryption provided by the development toolkit (i.e., akka) underlying remote message exchange that establishes a secure management connection between the central office and the remote PoPs. By testing the startup time and media channel configuration time (i.e., between end point phisical SBVT and through the emulated ROADM network) together with their components, in both distributions, at various network sizes (i.e., from

1 to 64 nodes), we are able to compare both distributions as well as asses the network scalability effects. Results demonstrate that our developed software is scalable by maintaining a relatively constant startup time for the networks tested (i.e., 1 to 64 nodes) of about 920 ms and 1150 ms for the centralized and distributed, respectively. Software scalability is also supported by the media channel setup time, which presents a modest log scale growth when increasing the number of nodes from 1 to 64. In both tests the centralized version presents lower startup and configuration times due to overhead in the distributed deployment caused by additional remote agent initialization procedures, encryption and java serialization.

As an additional contribution, the northbound interface of our domain controller was developed to support IETF flexi-grid specifications and offers a standard aggregated view of the network topology (i.e., IETF-network-topology) towards a higher hierarchical SDN or a multi-domain controller. Our controller successfully validates the coherent architecture of the IETF flexi-grid model and in the process, analyses and compares the most widely accepted and early specification information models (i.e. OpenROADM, OpenConfig, IETF flexi-grid), which represent one of the main interests of current research in the field of control and management of optical communications.

Chapter 3

Conclusions

3.1 Conclusion

This work presents a holistic approach to next generation fixed-mobile converged optical networks by tackling three fundamental segments: access/aggregation with a focus on architecture and functionality of NG-PoPs, intra and inter data center network solutions for high performance distributed NG-PoPs.

3.1.1 Next Generation Fixed-Mobile Converged Networks

Current mobile, fixed and wireless networks have evolved independently and as a consequence the network structure, devices, implementation, functionality are very different. This work pre This work regards FMC as one of the key strategies for next generation network architectures aiming to provide a better user experience while reducing overall cost and simplifying network operation and control through unified, access-agnostic, network functions. In order to achieve these objectives true convergence needs to be accomplished at both structural (i.e., shared network and infrastructure resources) and functional level (i.e., use of a generic universal set of functions to service all user connection types). Based on our analysis and on our integrated FMC demonstration, we envision the NG-PoP as the key element for both structural and functional convergence. On a structural level, it has been proven in **Papers A B** that WDM-PON systems in combination with our proposed low-latency cross connect can serve as a shared transparent and transport solution compliant even with the most rigid mobile baseband requirements (i.e., latency, jitter bandwidth). From a functional perspective, using SDN and NFV implementations inside the NG-POP, on-the-fly network and computational resource provisioning is accomplished for a series of services and universal functions (e.g., universal authentication, user mobility, function migration, content caching and serving etc.) as demonstrated in **Papers B** and **C**. In addition, SDN/NFV demonstrate an appealing multi-operator network sharing capability where multiple virtual instances of network functions inside the NG-POP, as well as virtual networks can be deployed over the same infrastructure.

3.1.2 Data Center Networks

Data center network display a rich and diversified design space. By evaluating the characteristics and performance for a set of highly interconnected topologies (i.e., torus, hypercube, jelly fish, fat tree) in **Papers C** and **E**, we infer that, even though the cost indications and performance metrics (i.e., bandwidth, latency, jitter, lost packets) are similar for all studied topologies with under 64 nodes, clear superior operations are displayed by the hypercube followed by the torus, on the downside of wiring complexity and scalability cost.

Focusing on control plane optimizations, we demonstrate in **Papers D** and **E** that, by employing an SDN-based centralized control plane even with a simple shortest path algorithm implementation, all topologies present significant performance improvements compared to conventional STP, reaching up to 45% more throughput in the case of hypercube for networks of 256 switches.

3.1.3 Control Plane for Inter-Data Center EONs

In **Paper G**, we proposed a scalable and modular flexi-grid optical domain controller based on Finite State Machines (FSMs). A key feature of our controller is represented by the deployment flexibility: centralized or distributed for added on-the-fly encrypted device configuration connections. Through our implementation of a flexi-grid optical domain controller, we validate the coherent architecture of the IETF flexi-grid model as well as demonstrates its practical application. A practical assessment of the model leads us to the conclusion that IETF flexi-grid YANG schema is mostly tailored towards network configuration however not towards monitoring. Therefore, additional extensions are required in order to support monitoring capabilities (e.g., pre-FEC BER, Q factor etc.) and allow dynamic reconfiguration of optical parameters based on changes in transmission factors.

By identifying and evaluating fundamental performance characteristics for our controller (i.e., startup time and media channel configuration time), together with their components, we demonstrate the scalability of our software. Both centralized and distributed approaches display relatively small influences on performance when startup and synchronization occurs. However, a modest logarithmic scale increase in delay for media channel setup is observed on network scaling. The FSM structure of the Network Driver allows for message pipelining and the multi-thread implementation presents good scaling capabilities. In both tests the centralized version presents lower startup and configuration times due to overhead in the distributed deployment caused by additional remote agent initialization procedures, encryption and java serialization.

3.2 Future Work

An overall extension of our work can be oriented towards developing a multi-domain control plane covering all three networking segments with the scope of providing an end-to-end service path orchestration spanning over the different domains. Communication and synchronization between the various domain controllers could be accomplished in two ways. One option would be to introduce a master orchestrator/controller in a hierarchical fashion with an overview of all interconnecting domains and capable of computing an end-to-end path across various domains. In this case redundancy and failure mechanisms should be addressed as the controller would be a potential single point of failure for all incorporated domains.

Another option could focus on a decentralized SDN architecture with horizontal communication protocols between domain controllers preserving domain privacy and preventing an overcomplicated multi-layer control mechanism. A decentralized approach could eliminate the disadvantage of a single point of failure brought by the previous alternative however would introduce the problem of data synchronization needed to avoid inconsistency in the network state.

Multiple research directions are also foreseen for each individual research topic presented in this thesis.

3.2.1 Next Generation Fixed-Mobile Converged Networks

Adoption of NFV and SDN technologies for the deployment of FMC networks may raise scalability and latency issues especially for more sensitive mobile traffic. Such concerns, could represent a good motivation for investigating further implications and considering software and hardware acceleration techniques for improvements.

Newer open reference implementations like OpenCORD (Central Office Re-architected as a Datacenter) aim at bringing datacenter economies and cloud agility to service providers for their residential, enterprise, and mobile customers. The potential for substantially decreasing the operational complexity of an FMC NG-POP using a CORD architecture can be twofold. Not only it provides automated operation but also deployment of preselected operating systems and software on bare-metal of servers and switches servicing various disaggregated technologies (e.g., vOLT, vBBU, vDOCSIS etc.) Furthermore, service function chaining can be achieved with the use of a Everything-as-a-Service Operating System (XOS) part of CORD capable of managing the SDN and NFV components together. Therefore, we believe that investigating a FMC design based on a CORD deployment could provide a valuable research direction.

3.2.2 Data Center Networks

Two research directions are foreseen in the area of data center networks. From a topological perspective, further investigations on developing more efficient and application specific interconnection graphs can serve as a starting point for future developments.

Furthermore, research focus can be directed towards new control plane optimizations like topology focused routing algorithms. As we've shown in this work, even a standard shortest path algorithm inside an SDN deployment can outperform a conventional STP installation. Additional performance improvements can be demonstrated by employing in SDN, multipath routing modules or developing custom algorithms for each topology and application running on top of them.

3.2.3 Control Plane for Inter-Data Center EONs

More in depth software performance characterization can be done with regard to various setups for example, accounting for actual controller placement in the management network and the delay incurred between the control and data plane. For a domain controller implementation, instead of Java, system programming languages could be used to improve performance and efficiency (e.g., ahead-of-time compilation languages like C, C++ etc.) In addition, development and evaluation of a path calculation element for a flexi-grid DWDM domain can be regarded as a possible future research direction.

SDN and SDO research is moving more and more towards integrated network management encompassing monitoring capabilities as well. Further extensions to the IETF flexi-grid model definition required by monitoring specific optical parameters (e.g., pre-FEC BER, Q factor etc.) represent a good target for future research. Providing support for such parameters, can lead to dynamic reconfiguration of optical transmission as part of new use cases in the area of network survivability and path restoration.

Paper A: Performance Evaluation of NETCONF-Based Low Latency Cross-Connect for 5G C-RAN Architectures

B. Andrus, A. Autenrieth, S. Pachnicke, S. Zou, J.J. V. Olmos and I. T. Monroy "Performance Evaluation of NETCONF-Based Low Latency Cross-Connect for 5G C-RAN Architectures," Accepted for publication in *International Conference on Transparent Optical Networks (ICTON)*, 2018, (*Invited*).

PERFORMANCE EVALUATION OF NETCONF-BASED LOW LATENCY CROSS-CONNECT FOR 5G C-RAN ARCHITECTURES

B. Andrus(1,2), A. Autenrieth(1), S. Pachnicke(3), S. Zou(1), J.J. V. Olmos(4), and I. T. Monroy(5)
(1)ADVA Optical Networking, Martinsried, Germany (bandrus@advaoptical.com)
(2)Technical University of Denmark, Department of Photonics Engineering, Lyngby, Denmark
(3)Kiel University (CAU), Kiel, Germany.
(4)Mellanox Technologies, Roskilde, Denmark
(5)Technical University of Eindhoven, Department of Electrical Engineering, Netherlands

ABSTRACT
The development of 5G wireless technology is in progress looking to cope with the increasing demands for high capacity, low latency, and ubiquitous mobile access. Cloud/Centralized Radio Access Networks (C-RAN) has been proposed as a promising approach to address the 5G benchmarks. C-RAN is a rising mobile network architecture based on the centralization and pooling of baseband processing elements, with the scope of increasing resource utilization efficiency and air-interface performance gains with fast scaling multi-cell coordination. In this paper we focus on proposing and evaluating a flexible low-latency cross-connect (XC) switch that allows a dynamic scheduling of networking resources between baseband units (BBUs) and remote radio heads (RRHs) in C-RAN deployments. On the one hand, we develop a control plane mechanism for manipulating the XC configuration and state data based on a custom YANG model and Network Configuration (NTCONF) protocol. Using a standard open NETCONF interface, automation of heterogeneous C-RAN resource assignment can be facilitated for 5G architectures. Secondly, we evaluate the performance of our XC in relation to the C-RAN stringent requirements and show that the latency and jitter introduced have negligible influence compared to radio interface limits. The impact of switching delay on the performance of a live system has yet to be tested however, a measured average switching time of 252 ms could still disrupt the ongoing connections.
Keywords — low latency XC, C-RAN, NETCONF, 5G.

1. INTRODUCTION

Given the strong growth trajectory of mobile data traffic [1], mobile networks need to evolve quickly in terms of coverage, capacity and new features, keeping a close eye on the stringent requirements concerning latency and data rates. In the view of ICT Metro-Haul project, new scenarios for future smart cities such as low latency real time object tracking and security are counting on strict network requirements provided by the upcoming 5G benchmarks. The advent of the Cloud Radio Access Network (C-RAN) [2] architecture seeks to meet the accelerating traffic demand while at the same time maintaining profitability and growth for the mobile operator. In addition, the expansion of mobile broadband internet gives a good opportunity for developing a cloud-based architecture, which enables a faster development process for new applications and services and which will support higher levels of energy efficiency [3].

In C-RAN, the baseband processing unit (BBU) along with its full functionality (baseband L1 as well as L2 and L3) is decoupled from the remote radio head (RRH) in an attempt to reach "full centralization". A new networking segment called fronthaul, emerges with the scope of connecting the remote radio units to the shared BBUs using high bandwidth, low latency optical access links. By breaking the static connection between the RRHs and BBUs new possibilities arise such as balancing traffic between different BBUs in accordance with mobile network load shifts, especially during the day. More precisely, the ability to counter the tidal effect caused by the migration of the business-area traffic towards the residential areas after the end of the working hours. Increasing fronthaul resilience is also a result of centralizing the protection switching mechanisms for the radio interfaces.

Fronthaul deployments will need to respect strict radio interface specifications like Common Public Radio Interface (CPRI) [4] in terms of latency, jitter and bandwidth. Network latency alone, can impact drastically user experience, especially when it comes to real time services therefore limitations start at 100 µs and often placed at 250 µs [5]. , For radio head synchronization and secure base station handover, jitter and its components have to be kept in tolerable ranges as indicated in Table 1. As baseband radio signal throughput is an order of magnitude higher than the IP payload carried, high bandwidth is required for the fronthaul.

Furthermore, fronthaul should also support load-balancing and switching capabilities [2]. Semi-static and adaptive BBU – RRH switching schemes for C-RAN architectures can lead to a decrease of 26% and 47% in the number of BBUs [6]. In order to take advantage of these features in a heterogeneous network and lead to optimization of resource usage, an open and standardized management and control of the elements is required. Network Configuration Protocol (NETCONF) [7] is emerging as a standardized protocol by the Internet Engineering Task Force (IETF), providing both control (e.g., data plane device configuration) and management functionalities by leveraging YANG based resource modelling [8] (language for network element description).

Table 1. Jitter upper threshold limits imposed by CPRI standards [4]; UI: Unit Interval

Jitter	Transmitter	Receiver
Total	0.35 UI	0.65 UI
Deterministic	0.17 UI	0.37 UI
deterministic and random (combined)	-	0.55 UI

In this work, we focus on introducing a cost effective solution for BBU – RRH switching in the form of an Optical-Electrical-Optical (OEO) XC with an open configuration interface. Chapter III of this work presents our NETCONF interface implementation based on a custom YANG model definition. Chapter IV details the performance evaluation with regard to the previously presented fronthaul requirements. Finally, we highlight our conclusion in Chapter V.

2. NETCONF Based Configuration Interface for Optical Cross-Connect Switch

NETCONF is a network configuration protocol that provides customizable control and management capabilities in a client-server architecture. Combinations of NETCONF and YANG (data modeling language), have been regarded with increasing interest by network operators because of the possibility to standardize common models for configuration and management data in a vendor-neutral way. Another advantage of NETCONF is the ability to provide a clear separation between state and configuration data.

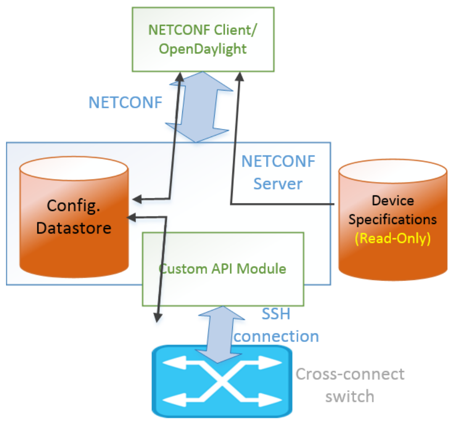


Figure 1. NETCONF client/server custom implementation.

In order to develop a NETCONF interface for our optical XC we used a NETCONF server implementation from Netopeer [9] that allowed us to define a new custom Application Programming Interface (API) tailored for our device. A detailed implementation and a testing setup can be observed in Figure 1. Northbound, the server exchanges XML based NETCONF messages (e.g., hello, edit-config, get-config etc.) with a client (e.g., OpenDaylight). Configuration data is stored in the server in an XML format file (i.e., datastore). Southbound, a custom API module has the role of translating the XML messages into custom device configuration messages.

On startup, the NETCONF server initiates the custom API module and triggers an SSH session to the XC. Once the session is established, the server queries the device for the current configuration state. Configuring actual device parameters is done through auto generated callback functions in the module. The functions are called based on the information present in the XML path language (XPath) of the NETCONF messages. For additional operations, Remote Procedure Calls (RPCs) are defined in the module. Relevant RPC examples in for the XC are reserved for actions like restart, enter or exit maintenance mode, power up/down.

Upon initial server-client connection, a handshake is initiated in which the client and server capabilities are exchanged and agreed upon. The NETCONF monitoring module sends information to the client about the datastores, locks and sessions as well as the YANG schema defining the device configuration. ODL also uses the downloaded schema to perform a local validation of any configuration request.

In order to define the semantics of operational and configuration data we have created a data model defining the capabilities of the XC with the use of YANG. The module defines the hierarchy of data used in NETCONF operations: configuration, state updates, RPCs. The data is organized in a tree structure with multiple named nodes that have either values or other child nodes attached. Our XC is modeled as a list of switching groups containing client (input) and network (output) ports complemented by a list of internal connections (e.g., connectivity matrix).

YANG definitions allows us to define special constraints on configuration data using “must” statements based on XPath expressions. Therefore, configuration errors can be caught in the client before issuing a NETCONF message to the server. For example, changing a port state on the switch or modifying an internal connection in a switch group is run through an additional

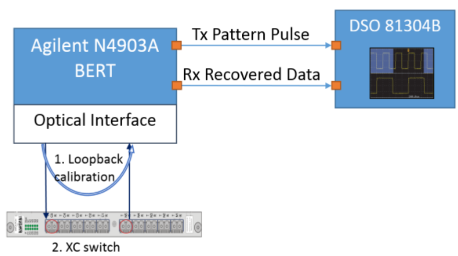


Figure 2. Setup for latency measurement of the cross-connect switch; 1st step calibration with a loopback fiber; 2nd step XC switch measurement. BERT: Bit Error Rate Tester; DSO: Digital Storage Oscilloscope.

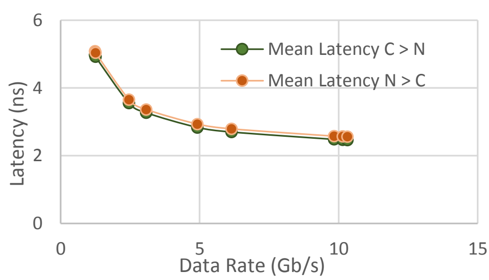


Figure 3. One way transmission latency: from Client to Network and from Network to Client measured at various CPRI line rates

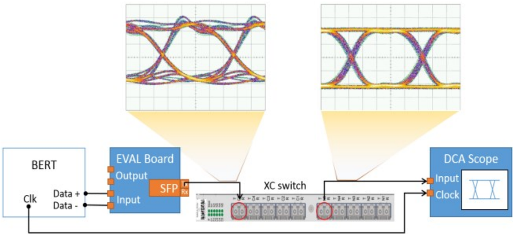


Figure 4. Jitter measurement setup: theeye diagrams shows the reshaping of the signal at the output of the XC switch corresponding to a line rate of 9830.4 Mb/s.

Table 2. Jitter measurements: periodic, random, deterministic and total jitter (UI) at the receiver and the transmitted side calculated for CPRI line rate options 4915.2, 6144 and 9830.4 Mb/s. Values are below CPRI specification thresholds.

Jitter (UI)		**4915.2 Mb/s**	**6144 Mb/s**	**9830.4 Mb/s**
Total	Trans.	0.171	0.214	0.302
	Rec.	0.147	0.16	0.311
Deterministic	Trans.	0.06	0.091	0.075
	Rec.	0.055	0.052	0.107
Random	Trans.	0.008	0.009	0.017
	Rec.	0.007	0.008	0.015
Periodic	Trans.	0	0.001	0.005
	Rec.	0.006	0.005	0.012

validation step that verifies, if the port-ID is present in the switch specifications or the input and the output port of the internal connection are in the same switching group. A snippet of the defined YANG model is shown in the APPENDIX.

3. Cross-Connect Performance Evaluation

By measuring the relevant performance metrics of our proposed cross-connect switch such as transmission delay, switching time, jitter at various throughputs, we aim to prove that a cost effective OEO XC complies with the C-RAN specifications.

A setup was implemented in order to test the device-induced latency. A Bit Error Rate Tester (BERT) Agilent N4903A was used as a data source due to its capability of allowing complete control over the data pattern. Inserting a custom and easy to identify sequence into the transmitted pattern with a recurring period of a magnitude order higher than the expected latency was transmitted and compared when received back. A high-speed digital oscilloscope is used to trigger a custom reference pattern for the delay measurement and record it on the receiving side as seen in Figure 2. The setup was calibrated using an optical loopback and measuring the total transmission latency for the link with a 2m long patch cord. Using the DSO, we measured the offset between the pattern pulse and the unique pattern identified in the recovered data.

As we can observe in Figure 3, latency varies between 2.5 – 5 ns. More exactly, as the data rate increases from 1.25 Gb/s to 10.1376 Gb/s the delay decreases from 5.04 ns to 2.57 ns. Moreover, the measured latency contains a fixed component of around 2.1 ns that is the result of the propagation delay through the fiber. Due to the Clock Data Recovery/Retiming function of the electrical cross-connect chip, an additional factor of approximatively 3.5 bit periods contributes to the final latency measurement. Therefore, as the data rate increases, the bit period decreases and the latency variable component decreases as well. This results in a better performance for higher data rates as shown in the graph in Figure 3.

Using the setup shown in Figure 4, jitter was measured at the input (transmitter side) and output of the XC. Corresponding eye-diagrams of the signal at the two measuring points are also displayed in the figure (@line rate 9.8304 Gb/s). We can observe the effect of the full 3R (Reamplification, Reshaping and Retiming) functionality of the XC switch comparing the shapes of the eye-diagrams at the receiver and the transmitter sides.

It is worth mentioning that different Small Form-Factor Pluggable (SFP) modules may vary the jitter slightly. In our measurements, we used 1310nm SFP+ on the transmitter side and SFP+ at 1574.4nm on the receiver side. The jitter results are documented in Table 2. The jitter sampling level used was set to 50%. The measured jitter values demonstrate the compliance of our tested XC switch with the rigorous fronthaul requirements (Table 1 – CPRI specifications) for all tested line rates. Moreover, the full 3R functionality of our XC has shown that in some cases it can even improve the jitter performance (e.g., total jitter at line rates 4915.2, 6144 Mb/s).

The combined output from two ports with a 50:50 coupler allowed us to monitor the signal variation at switchover as shown in Figure 5 and conclude that the normal signal level is restored after an average of 252 ms. The DCA Scope displaying the output was configured with a sampling interval of 1.5 ms.

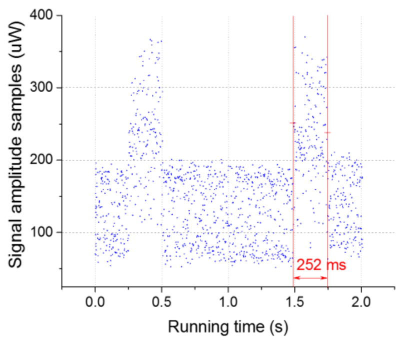


Figure 4. Switching delay results. Signal normal level is disrupted during port transitioning state.

4. Conclusion

Results in this paper can be regarded from a two-folded perspective. From a control plane aspect, we proposed and demonstrated a NETCONF protocol including a YANG model describing an optical XC. On the data plane side, we evaluated an XC and proved that a cost effective OEO implementation can fulfill the stringent requirements of radio interfaces such as CPRI in terms of latency and jitter at various bandwidths. Switchover between ports taking more than 200 ms will cause ongoing mobile connections to be dropped, however, we believe that the benefits of counteracting daily traffic tidal effects and improved energy efficiency could outweigh brief local service interruptions needed for dynamic fronthaul network reconfiguration or optimization.

ACKNOWLEDGMENT

This work was funded by the Marie Curie Project ABACUS, and by the ICT project Metro-Haul (G.A. 761727).

REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021, White Paper, 2017.
[2] China Mobile Research Institute, "C-RAN The Road Towards Green RAN", White Paper, 2011.
[3] A. Checko, et al, "Cloud RAN for Mobile Networks—A Technology Overview," Communications Surveys & Tutorials, IEEE, vol. 17, no. 1, pp. 405 - 426, 2015.
[4] Common Public Radio Interface (CPRI), Interface Specifications V6.0, 30/08/2013.
[5] H. J. Son, "Fronthaul size: calculation of maximum distance between RRH and BBU, 2014 available at: <http://www.netmanias.com/en/?m=view&id=blog_mbl&no=6276>.
[6] S. Namba, et al, "BBU-RRH switching schemes for centralized RAN," in 7th International ICST Conference on Communications and Networking, 2012.
[7] R. Enns, M. Bjorklund, J. Schoenwaelder, and A. Bierman, "Network configuration protocol (NETCONF)," IETF RFC 6241, June 2011.
[8] M. Bjorklund, "YANG - a data modeling language for the network configuration protocol (NETCONF)," IETF RFC 6020.
[9] NETCONF toolset Netopeer, found at: < https://github.com/CESNET/netopeer>.

APPENDIX

```
module optical-switch {
…
  container switch-specs {
    config false;
    ...
    list switch-groups {
      key switch-group-id;
      description "Switching groups in the device.";
      leaf switch-group-id{ type string; }
      list ports {
        key port-id;
        description "Ports in the switching group";
          leaf port-id { type string; }
          leaf port-type {
            type enumeration { enum CLIENT_PORT; enum NETWORK_PORT; }
          description "Predefined port types: Client or Network";  }}}}
    container switch-configuration{
      config true;
      list switch-groups{
        key switch-group-id;
        leaf switch-group-id{
          type string;
          must "count(/switch-specs/switch-groups[switch-group-id=current()]) = 1";
          }
         list ports {
          key port-id;
          leaf port-id {
            type uint32;
            must "count(/switch-specs/switch-groups/ports[port-id=current()]) = 1";
            }
          leaf port-state {
            type enumeration { enum ENABLED; enum DISABLED; }
            default ENABLED;
            description "State of a port";
          }
          leaf sw-group {
            type enumeration { enum 0x0100F110; enum 0x0200F110; }
            description "Switch-group belonging to"; }}
        list Internal-Connections {
          key in-port;
          unique out-port;
          description "Defining internal connections in the switch.";
          leaf IN-port{
            type uint32;
            must "count(/switch-specs/switch-groups/ports[port-id=current()]) = 1";
            must "/switch-specs/switch-groups/ports[port-id=current()]/port-type = 'CLIENT_PORT'" {
              error-message "Select client port"};
            description "IN connection port must be of Client type";            }
          leaf OUT-port{
            type uint32;
            must "count(/switch-specs/switch-groups/ports[port-id=current()]) = 1";
            must "/switch-specs/switch-groups/ports[port-id=current()]/port-type = 'NETWORK_PORT'"{
            error-message "Select network port"};
            description "OUT connection port must be of Network type";
    …
  rpc device-maintenance { description "Set device state to maintenance mode."; }
  rpc device-restart{ description "Send restart command."; }
```

Paper B: Demonstration of Next Generation Point of Presence for Fixed-Mobile Convergence

B. Andrus, R. Martinez, A. Autenrieth, M. Requena, R. Vilalta, B. Le Guyader, X. Grall, S. Gosselin, P. Olaszi, A. Pineda, A. Ladanyi, S. Zou, J. J .V. Olmos, I. T. Monroy "Demonstration of Next Generation Point of Presence for Fixed-Mobile Convergence," Submitted for publication to *International Journal On Advances in Networks and Services*, issn: 1942-2644, March 2018, (*Invited*).

Demonstration of Next Generation Point of Presence for Fixed-Mobile Convergence

B. Andrus[(1,9)], R. Martínez[(2)], A. Autenrieth[(1)], M. Requena[(2)], R. Vilalta[(2)], S. Zou[(1)], P. Olaszi[(3)]

[(1)]ADVA Optical Networking SE, Martinsried & Meiningen, Germany, (bandrus@advaoptical.com)

[(2)]CTTC Castelldefelds, Spain.

[(3)]AITIA International Zrt., Budapest, Hungary.

[(9)]DTU, Lyngby, Denmark.

B. Le Guyader[(4)], X. Grall[(4)], S. Gosselin[(4)], A. Pineda[(5)], A. Ladanyi[(6)] , J.J.V. Olmos[(7)], I.T. Monroy[(8)]

[(4)]Orange Labs, Lannion & Rennes, France.

[(5)]Fon Labs S.L., Getxo, Spain.

[(6)]BME University, Budapest, Hungary.

[(7)]Mellanox Technologies, Roskilde, Denmark.

[(8)]Technical University Eindhoven, Netherlands.

***Abstract*—Distributed data centers can benefit fixed and mobile services alike. Upcoming 5G technologies will force network operators to redesign current network infrastructures to deal with a high set of requirements (e.g., increased traffic load, reduced latency, improved cost- and energy-efficiency, etc.). In this regard, an appealing solution focuses on rolling out the so-called *Fixed Mobile Convergence* (FMC) in broadband networks. On the one hand, FMC aims at providing a shared infrastructure (i.e., transport solutions and common Points of Presence), and on the other hand, a set of universal functions and operations (i.e., authentication, accounting, path control and management, Caching, etc.) regardless of the access network type (fixed, mobile or Wi-Fi).**

In our vision, FMC is attained by developing a Next-Generation Point of Presence (NG-POP) based on characteristics of geographically distributed data centers. The NG-POP can be defined as the location for the common subscriber IP edge of fixed, Wi-Fi and mobile networks in FMC context. For a given area, user traffic flows from different access technologies are terminated within a single and common location, i.e. the NG-POP. Multiple NG-POP nodes can be deployed (in a distributed way) bordering access and aggregation network segments. In addition, within the NG-POP location, selected and common network functions and operations can be effectively hosted. To this end, we exploit the benefits of adopting both networking trends Software Defined Networking (SDN) and Network Function Virtualization (NFV).

In this work, we report on the successful validation of the devised and deployed SDN/NFV-based distributed NG-POP in the context of a FMC setup. The targeted FMC is attained by enabling within the NG-POP the support of heterogeneous (control and data plane) network functions for mobile core, Wi-Fi gateway and fixed services.

***Index Terms*—Universal Access Gateway; Fixed Mobile Convergence; Next Generation Point of Presence; SDN; NFV.**

I. INTRODUCTION

As our previous work [1] has shown that an SDN framework can significantly improve performance for high radix Data Center topologies such as hypercube, torus or jellyfish (e.g., as far as 45% more throughput per node), the complexity of scaling such networks proved to be an issue.

While the conventional trend regarding data centers is focusing on increasing their size and performance, an alternative approach turns towards a geographical distribution of data centers in key places in the network (e.g., NG-POPs), closer to the customers. Hosting business critical applications and IT infrastructure closer to the office location is preferred, in many situations [2], over the choice of a distant central location for reasons mainly related to lowered costs and latency.

Even though adoption of the afore mentioned topologies could not target conventional data centers, we could argue that a geographical distribution could alleviate the requirements of scaling internal networks to very large sizes. As such, distributed data centers could become a suitable deployment for high radix-networks therefore benefiting from the performance and resiliency advantages highlighted in [1].

Deployment of distributed data centers can provide added value not just for business applications. Mobile C-RAN architecture also seeks to apply data center technologies to allow for increased bandwidth, highly reliable, low latency interconnections in BBU pools. C-RAN imposes a set of stringent network requirements (in terms of increased traffic load, ultra-low latency, high availability, etc.) to support advanced services that network operators will need to provide. In this context, it is widely accepted that such infrastructures will deploy these services combining multiple resources such as networking (i.e., transmission, switching, etc.) and IT (i.e., computing, storage) [3].

In this work, we aim to introduce and demonstrate an NG-POP based on the characteristics of distributed Data Centers, that is able to service fixed and mobile access users alike and ultimately provide the basis for true Fixed-Mobile Convergence. With this objective in mind, in Chapter 2 we provide a brief introduction into the concept of FMC and the motivation behind a shared access infrastructure. The following chapter proposes the architecture for a distributed NG-POP (dNG-POP) detailing on its envisioned functionality. In Chapter 4, we describe the implementation of a full-scale demonstration setup representing a physical fixed mobile converged access network with a dNG-PoP at its center. Finally, we report the results from the first experimental demonstration and successful validation of the dNG-POP architecture for FMC access networks as part of the EU funded project, COMBO [4].

II. FIXED MOBILE CONVERGENCE (FMC)

The undergoing standardization process for next generation mobile networks (i.e., 5G) is expected to increase 1000 fold the wireless capacity introduced in 2010 and at the same time densely pack wireless links connecting up to 100 times more devices [5]. Such a prediction leads to assumptions on developing totally new backhaul and possibly adopt fronthaul technologies in order to cope with the increase.

Fixed access networks have been experiencing a significant increase for over two decades now (see Fig. 1), with no anticipated growth rate reduction. Ever changing multimedia and streaming services providing HD quality, or the newer UHD or even in 3D format, are some of the major bandwidth-hungry drivers today. Some of the current technologies trying to cope with these requirements, like the most often used DSL or cable (DOCSIS) over hybrid fiber-coaxial (HFC) infrastructure, offer connections of hundreds of Mb/s restricted however to a few hundred meters. In most cases already, fiber is deployed closer and closer to the end user surely leading to a fiber-to-the-home (FTTH) solution replacing copper.

Besides the increase of the overall throughput, as aforementioned, other 5G service and network demands (e.g., low latency, energy-efficiency, reduced CapEx and OpEx, etc.) need to be addressed by the network operator. These requirements are handled from and end-to-end perspective covering several network segments and multiple technologies (i.e., mobile, Wi-Fi, packet and optical switching, etc.). As a consequence, this end-to-end vision significantly challenges network operators which aim at rolling out targeted 5G networks in a cost-efficient manner to maintain their competitiveness.

Bearing the above aim into mind, an appealing approach gaining momentum to deploy cost-efficient 5G network is based on integrating and merging traditional independent network

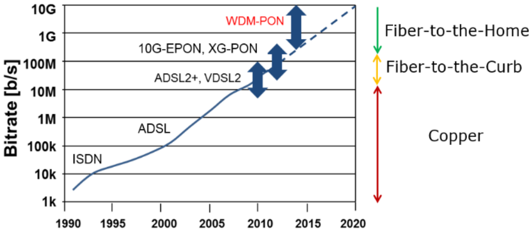


Fig. 2. Evolution of access technologies [6]

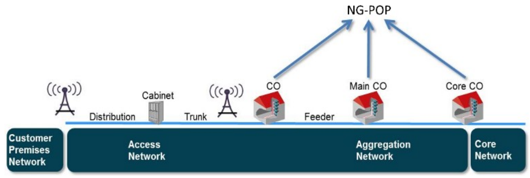


Fig. 1. Reference locations for Fixed and Mobile Network Integration NG-POP (CO: Central Office)

infrastructures for fixed and mobile traffic services into a common network and control functions. This is referred to as Fixed Mobile Convergence (FMC) and currently is envisioned within the 5G networks roadmap [7].

We have previously shown in [8] that a FMC architecture should target solutions for cost-efficient FMC from a twofold perspective: structural and functional convergence. The former focuses on sharing and unifying equipment/technologies (at both access and aggregation network segments) to transport seamlessly both fixed and mobile traffic (e.g., via a WDM-PON infrastructure). The latter seeks for common set of control functions (e.g., unified control and management, AAA, etc.) to handle any access service type. Both objectives can be achieved by deploying the network architectures based on NG-POP. NG-POP is defined as a network location featuring a number of control and data plane FMC-driven capabilities, e.g. unified IP layer gateway (Universal Access Gateway - UAG), BBU hostel (for CRAN applications), caching server for content delivery networks, unified authentication, etc. When NG-POPs are distributed in a large number of locations, close to the user, they can also host Access Node functions such as OLTs or BBU hostel (for CRAN applications).

Two NG-POP scenarios are foreseen (see Fig. 2): i) distributed, NG-POPs deployed in a large number of locations, e.g., between access and aggregation networks (i.e., Central Office –CO – or main CO – on a higher aggregation level than a regular CO however still not connected directly to the network core); ii) centralized, NG-POPs placed in a small number of locations, e.g., between the aggregation and core networks;

For both, the implementation leverages the benefits of current networking trends: centralized Software Defined Networking (SDN) control, and instantiation of Virtualized Network Functions (VNF) in commodity servers (applying the Network Function Virtualization -NFV concept). Both distributed and centralised architectures are valid approaches for rolling out the NG-POP solution. In this work, we will focus on the demonstration of dNG-POP deployment.

A. FMC Shared Access Network

As structural convergence in the access and aggregation segment requires not only more capacity, but also extensive reach and potential transparency [9]. To this end, the wavelength division multiplexing passive optical network (WDM-PON) is adopted to handle fixed access, Wi-Fi backhaul, and mobile backhaul/fronthaul, as shown in Fig. 3. The WDM-PON technology is able to cope with the high capacity demand of expected 5G fixed and mobile advanced services, and also guarantee a smooth evolution of the legacy access networks. In our demonstration, two different types of WDM-PON are explored for different use cases, namely, a Wavelength-Selective

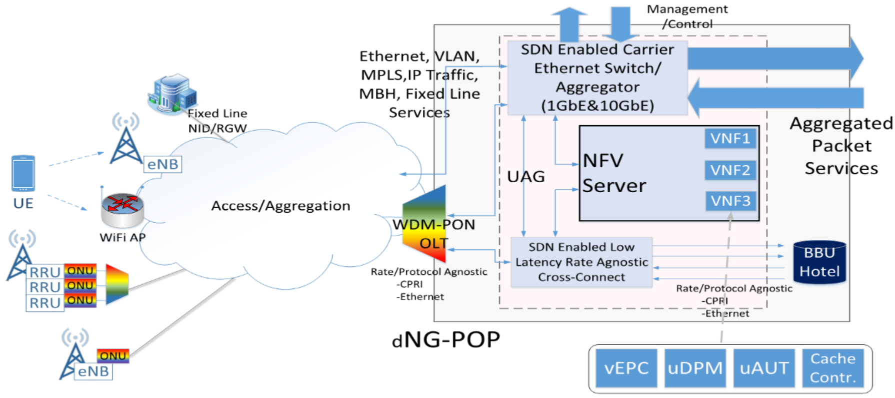


Fig. 3. Fixed Mobile Convergence system architecture with a shared Access/Aggregation network converging in a NG-POP. (UE: user equipment; UAG: Universal Access Gateway; vEPC: Virtual Evolved Packet Core; uAUT: Universal Authentication; uDPM: Universal Data Path Management)

(WS) WDM-PON and a Wavelength-Routed (WR) WDM-PON [10]. The WS-WDM-PON is feasible to be upgraded from the legacy power splitter based optical distribution network (ODN), by using both tunable transmitter and receiving filter at the optical network unit (ONU). Alternatively, the WR-WDM-PON adopts a novel full C-band tunable laser at the ONU, and a cyclic WDM multiplexer/de-multiplexer at the remote node to route a single wavelength to the corresponding ONU. Such a WDM-PON solution is especially suited to the aforementioned requirements of a converged infrastructure with regard to the bandwidth × reach product (e.g., bandwidth of up to 10 Gb/s per wavelength and reach of > 50 km), which are not supported by today's existing WDM-PON approaches. xfhfgh

B. *Distributed NG-POP Architecture: Main Features*

A feasible implementation of the dNG-POP architecture targeting FMC objectives is depicted in Fig. 3. The main components (building blocks) are highlighting along with the access network infrastructure used to connect transparently various client access point technologies to the dNG-POP entity. The main physical components, upon which the dNG-POP is built are:

- An NFV server
- A low-latency cross-connect
- A provider Ethernet switch/aggregator

The role of the NFV server focuses on providing support for the functional convergence. That is, the aforementioned VNFs are hosted onto an off-the-shelf server running a customized cloud environment. Breaking the static one-to-one BBU-RRH implemented by BBU hoteling is realized through the low latency cross-connect. Such a cross-connect complies with the rigorous latency and jitter timing requirements of the Common Public Radio Interface (CPRI) [11] between the RRH and the BBU agreed on by major system vendors. Aggregating the various user connections from a number of access devices onto higher line rate links is done by the provider Ethernet switch/aggregator. I addition, it is also responsible for identifying the various access channel types and isolating them into VLANs.

The functional role of the dNG-POP can be concentrated in the scope of a UAG seeking to providing control over all user sessions by taking advantage of resources available within each access network. In our implementation, the NFV server represents the unique point in the network where data flows of any user coming in from any type of network can be accessed by the control plane.

Moreover, the need to have a centralized, intelligent network entity that can dynamically allocate and reconfigure data paths converging inside the UAG has led to the adoption of an SDN approach. The control plane functionality (network element configuration) of the Ethernet, the NFV server internal network and the cross-connect switches is handled by an SDN controller (relying on the OpenDayLight implementation) via OpenFlow [12] and NETCONF interface [13], respectively.

The devised dNG-POP architecture is targeting a pool of use cases which are executed to validate a number of different network functionalities running on the NFV server such as the uAUT, uDPM, vEPC and vCache. These functional convergent-oriented operations cover both control and data plane functionalities and are discussed in the following sections.

1) *Universal Subscriber and User AUThentication (uAUT)*

Resource access control is one of the most important functions in a network regardless of the access technology. Indeed, there are specific authentication techniques for each network type. In a standard scenario with multiple access network types, an operator needs to assign credentials for users in each network and solve the authentication and the accounting in each of them independently. This is not efficient since separated and isolated mechanisms and databases need to be maintained by the network operator which increases the complexity of the whole system.

The proposed Universal Authentication (uAUT) system is a basic function of the UAG that offers support to all additional control

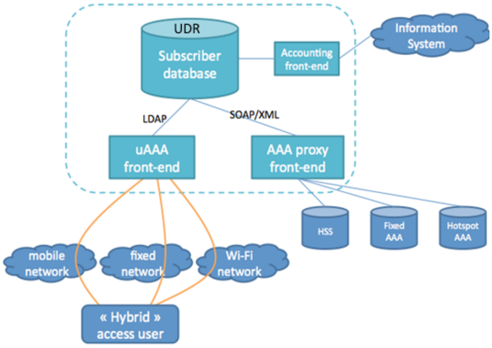


Fig. 4. uAUT architecture (AAA: Authentication Authorization Accounting; HSS: Home Subscriber Server; UDR: User Data Repository)

plane functions of the NG-POP. Its main task is to provide authentication authorization and accounting (AAA) to users associated with all access networks serviced by the NG-POP. Its usage is mainly restricted to the initial phase of a service setup (e.g. provisioning policies at the network attachment) and accounting of the service delivery for billing and auditing purposes. uAUT serves as a unique contact point within the UAG for all subscriber data and authentication related functions, regardless of the access type employed.

The proposed uAUT architecture, maintaining legacy compatibility, is presented in Fig. 4. The architecture is based on the User Data Convergence concept [14] and supports a layered architecture, separating the user data storage from the application logic. The view is extended by storing the user data in a unique User Data Repository (UDR) which provides a unified view for subscriber management to the information system (billing, accounting, statistics etc.). Dedicated entities handling application logic, named front-ends (FE) represent the links between fixed/mobile network services and the user database. Examples of network services that need to access user data include: mobile Home Subscriber Server (HSS), Wi-Fi hotspot AAA, broadband AAA, Policy Control and Relay function etc.

The UDR hosted by the uAUT server allows the service provider to identify all user connecting to any access type. By mapping to the correct profile, users can receive access to the converged services such as unified accounting, seamless authentication to application platforms (e.g., IPTV, VoD) and Over-The-Top (OTT) partners.

From the user point of view, the uAUT functional block provides a common subscriber authentication platform allowing the UE to login from both Wi-Fi and mobile networks. This is accomplished by using the same credentials stored in the SIM card. In the experimental demonstration, a vEPC instance running on the NFV server stores the authentication key in the EPC Home Subscriber Server – HSS function. The so-called hybrid access is achieved by accessing the mobile credentials through the common AAA proxy front-end. Further in-depth technical details on the proposed hybrid access architecture have already been presented in [15] and demonstration results are described in Section IV.

2) Universal Data Path Management (uDPM)

High proliferation of mobile broadband communications targeting 5G networks is expected. An important objective of FMC which addresses this concern is mobile traffic offloading and handover. Implications like metro and core offload are also foreseen. Allowing users to roam between fixed/mobile/Wi-Fi networks and transport traffic via several types of interface requires a converged subscriber and session management as well as an advanced interface selection and route control. This set of functional blocks fulfil the scope of a Universal Data Path Management (uDPM).The uDPM is the main entity of the data path control functions performed within the UAG responsible for providing the UE session continuity.

From a FMC user's perspective which is connected to various access points, numerous data paths can be used concurrently for increased Quality of Experience (QoE) or as backup for seamless handover. Multipath TCP (MPTCP) is a TCP extension [17] making use of end-to-end path diversity and maintaining backward compatibility. Protocol operation establishes several different TCP subflows (e.g. remote/local IP and port) for concurrent data traffic managed by a main MPTCP connection (between two end points). In our scenario, the use of MPTCP enabled UEs and content servers mitigates connection interruptions at network access switch over.

Moreover, the uDPM architecture consists of several interconnected and dedicated functions as shown in Fig. 5. These functions control and handle session mappings of each individual UE to multiple data paths. A monitoring function that collects user and network state information can create a session event relative to a UEs activity (e.g. application launch request, interface change request, data forwarding process etc.) and trigger the uDPM functional block. Session event notifications include signal degradation detected by the UE or network, discovery of a new access point, applying a network policy or a subscriber profile, etc.

A Decision Engine (DE), being in part under the operator's control, uses an algorithm to check network operator policies and

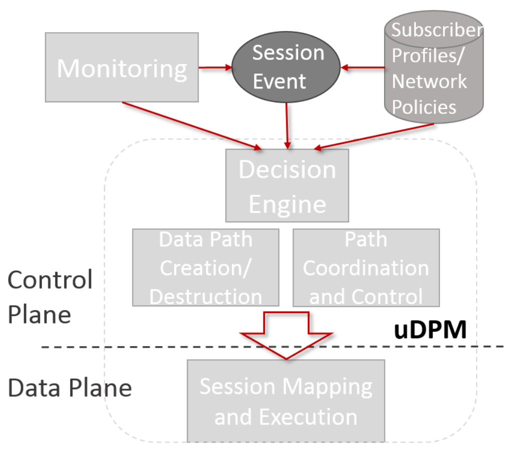


Fig. 5. Universal Data Path Management (uDPM) functional blocks

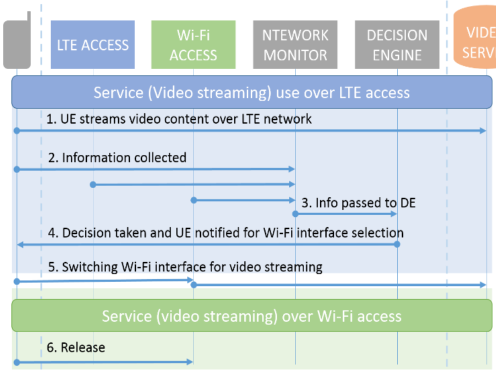


Fig. 6. Decision Engine Operation workflow example

subscriber's profile rules. The algorithm relies on multi-criteria decision making required by processing multiple rule categories. The output of the DE can involve creating/destroying data paths (data path creation/destruction block) or seamless network handover in terms of session continuity (path coordination and control).

When a session is based on multiple paths, there is a coordination requirement of those data paths which, within the uDPM architecture, is conducted by the Path Coordination and Control element. This element ensures session continuity where data traffic is transferred correctly and effectively over a number of established paths.

Session mapping execution, as part of the data plane, applies session mapping decisions taken by the DE by relying on the control of both "path creation/destruction" and "path coordination and control". Session packets are forwarded or filtered on the data path and subflows are merged in MPTCP connections.

The DE algorithm can take into account different sources of information for its internal computations, like: network related information (Wi-Fi APs and mobile BS location, traffic load, energy consumption etc.), subscriber information (profiles or QoS classes) or content information (cached content). A workflow exemplifying the Decision Engine mode of operation is displayed in Fig. 6. In the first step, a UE requests the stream of an internet video over an LTE network. The network monitor function (polling UE, LTE and Wi-Fi interface and network status) feeds the decision engine algorithm. Evaluating the input information according to its preconfigured targets (e.g. cost and bandwidth optimization), the DE decides to switch the streaming session from LTE (lower bandwidth and higher cost) to Wi-Fi (higher bandwidth and lower cost). This is done by notifying the UE to switch the active connection from its mobile data to its Wi-Fi interface. Finally, the UE streams the video content over Wi-Fi.

A more detailed description of the uDPM technical implementation has been presented in [14] and demonstration results for relevant use cases are shown in Section IV.

3) Content Distribution System

Content distribution techniques aim at reducing the redundant traffic in the network and improving quality of delivered services. A converged Content Delivery Network (CDN), in the context of FMC, can achieve this goal with better reliability, scalability and performance.

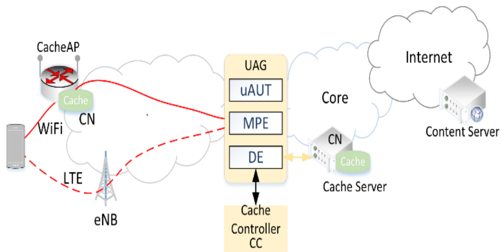


Fig. 8. Converged content delivery system

Caching efficiency is directly proportional to the user density on a network segment. The less population, the less useful caching is. In this regard, research studies have shown that in fiber access networks (30 thousand UEs) as well as in xDSL infrastructures more than 30% of the traffic can be reduced due to the fact that almost half of the requests are cacheable [16] and [17]. The situation is somewhat different in the case of a mobile networks. According to the studies made in [18], caching at the base station (eNB) or at home gateway does not bring improvements. However, implementing a content delivery solution in a converged network the advantages can be multiplied with a collaborative caching algorithm. Measurements performed in [19] support the cooperation between telecom and CDN providers. Such a collaboration leads to an additional traffic decrease of 12 – 20% if collaborative content caches located in NG-POPs are implemented.

A content delivery system is developed (Fig. 8) comprising of a Cache Node (CN) and a Cache Controller (CC). In this custom implementation, the CN, located in the home gateway in the form of a Cache Access Point (CacheAP: a wireless access point with caching functionality) but also in the NG-POP, executes the caching and prefetching. The virtual Cache Controller (vCache) installed in the NFV server (within the NG-POP) is responsible for managing the caching functionality and providing Caching-as-a-Service to content service providers.

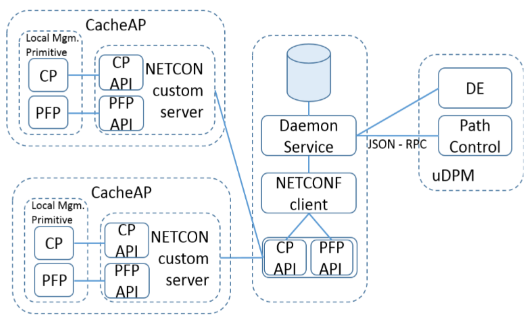


Fig. 7. Caching/Prefetching system architecture [20] (pp. 64);

The content delivery system relies on the uAUT functionality even though in the first authentication phase (when the user authenticates in the network) the uAUT does not influence the delivery service execution. However, extending the authentication from the network level to the service level requires interaction between uAUT and the content delivery system. The goal is to provide transparent service delivery using a unified authentication process.

The caching system architecture explained in detail in [20] (pp. 64) and used in this demonstration (Fig. 7) is divided in three main components:

- A CN in the form of a CacheAP based on a custom NETCONF [11] server implementation and a local management primitive that manages the local caching/prefetching actions;
- The CC composed of a daemon service that exchanges JSON based Remote Procedure Calls (RPCs) with the uDPM module; a NETCONF client for communicating and managing the CN; and a data base that stores: CN config, user requested content and content already cached in the CN;
- uDPM module described in the previous section with a DE and Path Control functions;

There is a tight dependency between the content caching system and the uDPM as seen in Fig. 7. When the number of hits on a specific content increases over a predefined limit (i.e., threshold), the content provider can trigger the caching procedure indicating the content (stream URL). The Decision Engine provides the needed resource information regarding UE location (client ID and IP, cacheAP IP) and network performance. Such information is the input required by the cache controller to make an optimal caching decision and prefetching the contents in a CN (closer) to the UE.

Further details on the caching-system implementation can be found in [20] (pp. 63 – 66).

III. FMC Demonstration Setup

The FMC setup (shown in Fig. 9) used for the final demonstration was deployed aiming at validating the feasibility and evaluating the efficiency of the proposed dNG-POP concept and its developed features.

The main elements and technologies constituting our experimental setup are:

- A WR and a WS WDM-PON systems enabling the shared access network infrastructure to carry transparently Wi-Fi, mobile (LTE), CPRI and fixed subscriber traffic.
- dNG-POP (located at the main CO) implements the set of control and data plane functionalities needed by the common subscriber IP edge for all traffic types (i.e., fixed, mobile and Wi-Fi).
- The transmission is made over 18 km span of Lannion city fiber ring showcasing the capability of real field deployment;
- Heterogeneous endpoint access equipment formed by LTE base stations (eNBs), Wi-Fi access points (APs), cache AP.
- User test devices (e.g., laptops and smartphones) located in the customer premises area;

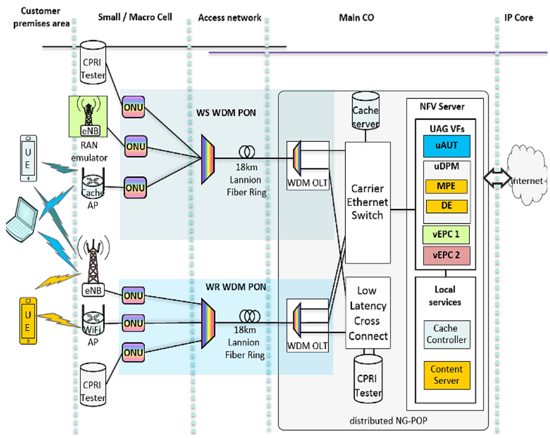


Fig. 9. Demonstration setup overview

A. Shared access network: WDM-PON

Both WR and WS WDM-PON systems employed in this demonstration have been evaluated as suitable candidates for a proposed convergent access and aggregation network architecture by providing high reach, high data rates and low latency and jitter.

Tests carried out on the experimental WR-WDM-PON system, supporting the functional demonstration, show compliance with the CPRI standard even at the highest line rates (9.83 Gb/s). We have shown previously, [21], that the maximum jitter specifications on the receiver and the transmitter side are fulfilled. Measurements performed also showed that the system induced latency is as low as 130 ns which is equivalent signal propagation through only 26m of optical fiber. Evaluation of our developed WS-WDM-PON system in [20] (pp.88 – 90) using an Integris Mobile Access Network Performance Tester showed a consistent latency of 91.59 ns and a BER of $2.3{\cdot}10^{-15}$ measured for 2.45 Gb/s data rate. System attributes of our developed WDM-PON systems are compared in Table 1.

In our current demonstration, the wavelengths were transmitted over an 18km fiber ring deployed in Lannion (France) testing the capabilities in real field deployments.

B. dNG-POP setup

At the core of the dNG-POP, the NFV server is built on an OpenStack cloud system. Features like automated configuration and on-demand resource deployment make OpenStack an ideal platform for our demonstration. The support for allocating various computing and networking resources for each targeted functionality, isolating them into individual projects (e.g., EPC, uAUT and uDPM) is

Table 1. WDM-PON systems performance.

WDM-PON	Latency	BER
WR	130 ns	10^{-12}
WS	91,59 ns	$2.3{\cdot}10^{-15}$

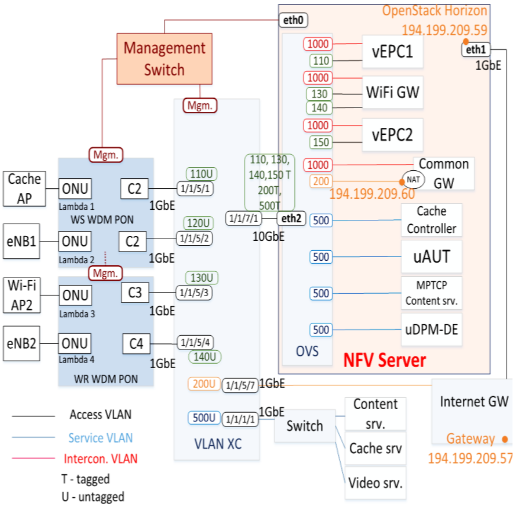


Fig. 10. Network control plane overview with VLAN setup assignment

perfectly tailored for our setup. In this scenario, we observe that multiple (and independent) instances of the same functionality could be instantiated within the NFV server as long as sufficient (computing) resources are available. This provides the dNG-POP with the capability of supporting multi-operator network function instantiation. This means that different operators may have their own network functions deployed within the same physical host (NFV server) but without having visibility of other operator's network functions. To show this, two instances of the mobile core (i.e., vEPC1 and vEPC2 in Fig. 9) were deployed as VNFs onto the NFV server.

The network control plane overview with VLAN assignment in Fig. 10 shows the seamless synchronization achieved between the OpenStack cloud environment and the hardware infrastructure. In this sense, the UAG's Carrier Ethernet switch (ADVA FSP 150EG-X) acts as a VLAN cross-connect, isolating individual connection types into separate VLANs with unique IDs. In our setup, the access channels are numbered from VLAN 110 through 160. More exactly, VLANs 110 and 150 are used for identifying the LTE S1 interface control and data traffic backhauled from the two eNBs over the WDM-PON to the corresponding vEPCs. Multiple wireless APs destined for individual test cases are mapped with VLANs 120 through 140, and the cacheAP is on VLAN 160.

The stand-alone local services (e.g., video, content, caching servers) and testers are grouped into a common service VLAN 500. The interconnection channel supporting network function chaining is handled in VLAN 1000. On this VLAN, installed on the NFV server, a Common Gateway handles the network address translation (NAT) for all VNFs providing them with Internet access.

Maintaining a Layer 2 network setup continuity from the physical infrastructure inside the NFV server was accomplished by configuring OpenStack to have access to the provider network through its SDN enabled Open Virtual Switches – OVS. A 10 GbE

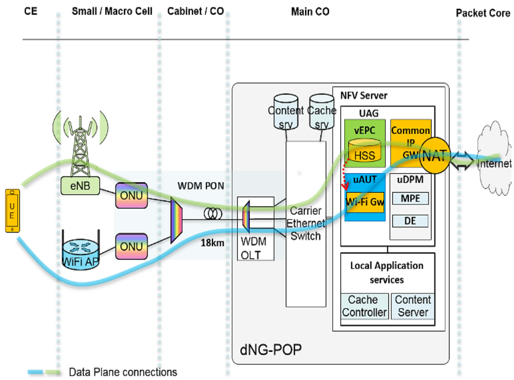


Fig. 11. uAUT functional demo overview

optical line card was set up to connect to the Carrier Ethernet switch and the NFV server to effectively handle the user data plane traffic for all targeted test cases. Even though in our demonstration we used a manual configuration of VLANs, in a real live deployment an automated SDN controller assignment is expected.

IV. Experimental Validation

The demonstration of the proposed and implemented capabilities of universal authentication, user mobility and content caching solution is carried out by individual test cases which are executed over the setup detailed in the previous section.

A. Universal Authentication (uAUT)

Two access points are set up to provide simultaneous network connectivity to UEs through wireless and LTE access network types (Fig. 11). Both the Wi-Fi AP and the eNB are connected through the same access network infrastructure (i.e. WDM-PON) to the UAG's NFV server where the vEPC Wi-Fi Gw and uAUT are instantiated as individual VNFs.

A UE is used to test the authentication by presenting a SIM card with the common set of credentials. Using the SIM, the UE can transparently and seamlessly authenticate in both access technologies (i.e., mobile and Wi-Fi).

The first step in the functional use case is the validation and evaluation of user authentication in the LTE network. In this scenario, the user request is sent from the eNB to the vEPC over the S1 interface. The entire LTE user-attach procedure (Fig. 12) was measured (on average) at 650 ms, including the user authentication phase which took around 279 ms. Measurements were performed with the use of Wireshark, a network protocol analyzer.

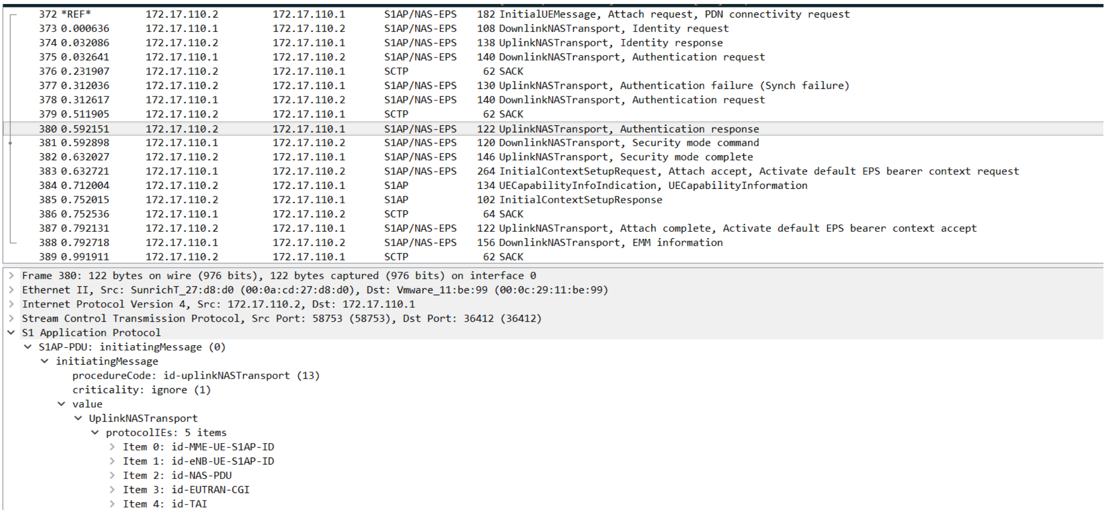


Fig. 12. LTE attach procedure (Wireshark capture trace) containing the authentication phase (LTE S1 interface: used for communication between eNB and EPC)

In the second step, the user performs a switchover to the Wi-Fi network. Upon the users' authentication request, the Wi-Fi AP is configured to send the connection request to the uAUT server residing on the NFV server. The request is processed by the uAUT which compares the credentials received from the user with the credentials stored in the HSS element of the vEPC VNF. Retrieving the credentials from the HSS was accomplished by implementing an extensible authentication protocol framework for UMTS (EAP-AKA). Measurements showed that authentication phase took 10ms over Wi-Fi.

B. *User mobility demonstration*

The second scenario executed reports the offloading and handover process especially between mobile and wireless networks. This allows a UE to efficiently use the network resources. For the user mobility demonstration, two UEs, the LTE eNB and two Wi-Fi APs (TP-Link TL-WR1043ND), the uDPM VNF have been employed as well as a MPTCP Content Server positioned in the Local Services area of the dNG-POP (Fig. 14).

Using a custom API, the uDPM-DE provides information to the UE regarding the access method selection. In this context, a set of feasible scenarios is executed outlining automatic and even seamless handover process. The lack of service interruption during the handover was ensured by the use of MP-TCP function in the NFV Server.

Three use cases have been conducted in order to demonstrate and evaluate the efficiency of the uDPM functionality:

- a plain Wi-Fi to Wi-Fi handover corresponding to a use case in which a UE will be transferred to another AP when

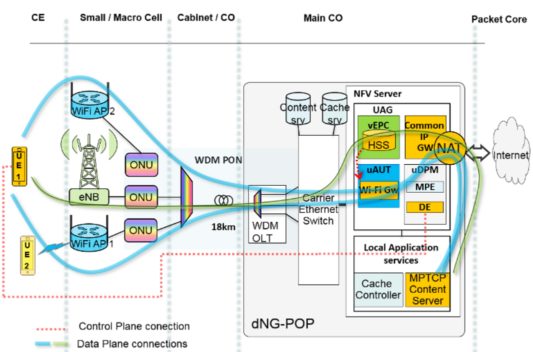


Fig. 14. User Mobility demonstration overview

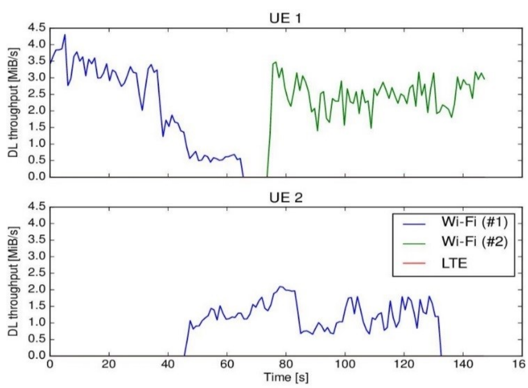


Fig. 13. Wi-Fi to Wi-Fi handover – connection gap visible

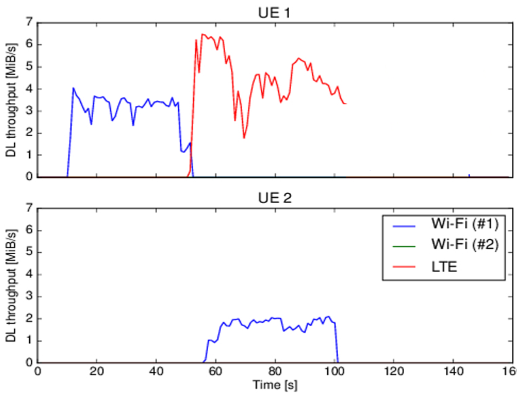


Fig. 15. Wi-Fi to LTE handover – no connection interruption

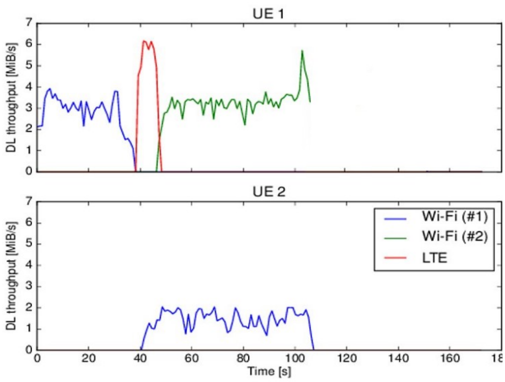


Fig. 17. Wi-Fi to LTE to Wi-Fi handover – no connection interruption

the current wireless link is saturated;

- a Wi-Fi to LTE handover corresponding to a use case where a UEs' ongoing connection will be switched from the current saturated AP over to LTE;
- A Wi-Fi_1 to LTE to Wi-Fi_2 handover. This use case is an improvement on the first test where there is a gap in the connection switch which is now filled by a transient LTE connection.

Firstly, a Wi-Fi roaming from one AP to another was tested. A connection was established between the UE1 and the content server by requesting a video stream over Wi-Fi_1. Soon after, a second UE (UE2) connected to the same AP starts a download and saturates the link.

The DE that is monitoring the network state, triggers the handover of UE1 to the available Wi-Fi2 AP in order to offload the former wireless link. As observed in Fig. 13, there is a connection gap of about 10s during the handover which is the result of using a single wireless interface on the client device.

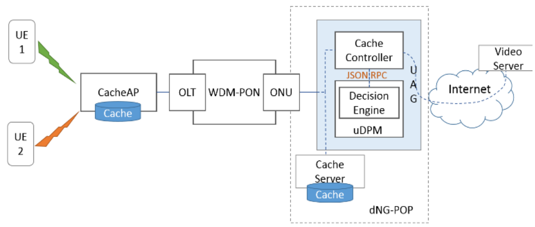


Fig. 16. First caching test case demonstration: UE 2 requests the same video content as UE 1 after it was cached in the CacheAP;

The second test case shows a Wi-Fi to LTE handover triggered by the DE in similar circumstances as the previous one. When the Wi-Fi link is saturated, the DE triggers the switchover to the available LTE interface. In Fig. 15 we can observe the seamless transition between the two networks. The connection is uninterrupted because the UE can be connected simultaneously to both networks.

The last test case provides a solution to bridge the connection gap between inter Wi-Fi handover by transiently switching from Wi-Fi_1 to LTE then over to Wi-Fi_2. We notice in Fig. 17 that employing this method, video streaming was uninterrupted. We also observe a small overlapping traffic pattern in the case of LTE to Wi-Fi_2 handover due to packets duplication over the two MPTCP subflows. However, data is correctly reassembled by the master MPTCP session.

C. *Content delivery service (content caching)*

For the caching demonstration, an SDN-based Cache Controller VNF (vCache) is instantiated on the NFV server. It decides where (e.g, at either access network device – CacheAP- or the dNG-POP) to cache or prefetch the content. The CC and uDPM-DE are coordinated to instruct any UE to connect to a different CacheAP as long as the QoS is degraded due to congestion in the CacheAP node or if a better connection is available.

For the test case two UEs and two CacheAPs (mobile AP with caching and routing capabilities) have been employed. Two test cases have been performed, one highlighting the caching ability and one focusing on the prefetching execution.

In the first caching test case (Fig. 16), A UE streams a video from

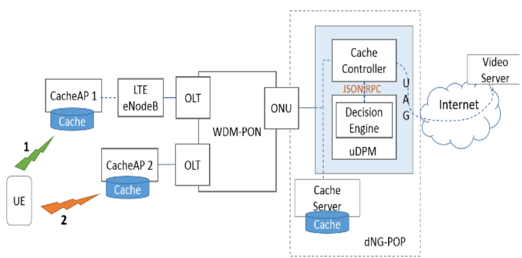


Fig. 18. Second caching test case demonstration: CC prefetches the video content on second CacheAP when UE switches to another network.

the internet (YouTube) with a bandwidth requirement higher than the network bandwidth allocated. Traffic Control (TC), a linux network utility used for traffic shaping, was used to set video bandwidth limitations. The QoS of the video is visibly degraded (long startup delay, frequent interruptions). After the request, the content is cached automatically in the CacheAP. When the second user (UE2) requests the same content, the video is delivered from the CacheAP and the observed quality is greatly improved (no more buffering timeout periods).

The second prefetching test (Fig. 18) make use of two CacheAPs and one UE. The Cache Controller holds the responsibility of making an optimal prefetching decision based on user profile information (user ID, URL of video played) as well as network status (network address, current AP, destination AP) received from the DE. Once the CC has computed the caching location (CacheAp address) the decision is sent back to the DE which will handle the interface selection mechanism for the end user. The trigger of the prefetching is a UE handover from the first CacheAP (LTE network) to the second one (fixed line). This situation can correspond to multiple scenarios (e.g. current network saturation, a user arriving home and switching to the local network etc.). The switchover commanded by the DE is also passed to the CC along with source and destination AP. The CC then retrieves, from the user profile, the video URL and sends it along with the fetch command to the destination AP. By the time the UE has switched interfaces, the video has already started being cached in the second AP.

D. *Client channel bandwidth testing*

In order to consolidate the demonstration of the fully integrated setup and evaluate its impact on the end user, we tested the TCP bandwidth from each network access type (Wi-Fi, LTE, fixed). As seen in Fig. 20, TCP bandwidth and latency were measured between a client and a common IP core gateway on the NFV server regarded as the reference point. Iperf, an IP network measurement tool and ping were used to measure the performance on each channel consecutively. Relevant settings like Maximum Transmission Unit (MTU) with a default value of 1500 MTU and test intervals of 10 s were configured. A test report capture of an LTE network is present in Fig. 19.

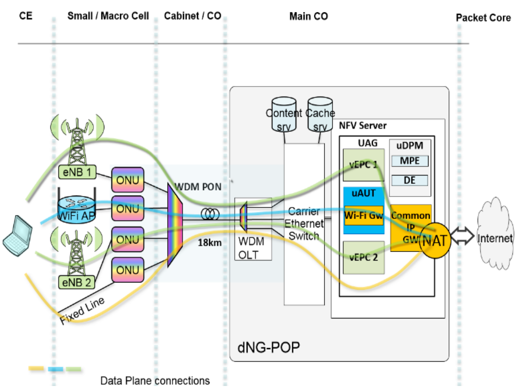
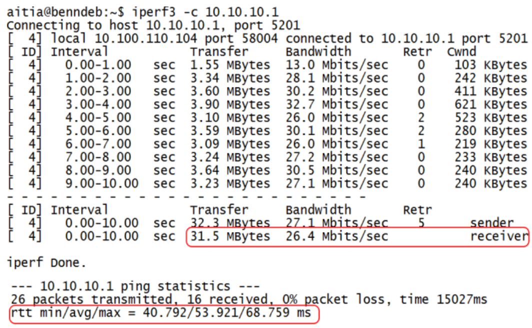


Fig. 20. Overview of client channel bandwidth test



```
aitia@benndeb:~$ iperf3 -c 10.10.10.1
Connecting to host 10.10.10.1, port 5201
[  4] local 10.100.110.104 port 58004 connected to 10.10.10.1 port 5201
[ ID] Interval           Transfer     Bandwidth       Retr  Cwnd
[  4]   0.00-1.00   sec  1.55 MBytes  13.0 Mbits/sec    0    103 KBytes
[  4]   1.00-2.00   sec  3.34 MBytes  28.1 Mbits/sec    0    242 KBytes
[  4]   2.00-3.00   sec  3.60 MBytes  30.2 Mbits/sec    0    411 KBytes
[  4]   3.00-4.00   sec  3.90 MBytes  32.7 Mbits/sec    0    621 KBytes
[  4]   4.00-5.00   sec  3.10 MBytes  26.0 Mbits/sec    2    523 KBytes
[  4]   5.00-6.00   sec  3.59 MBytes  30.1 Mbits/sec    2    280 KBytes
[  4]   6.00-7.00   sec  3.09 MBytes  26.0 Mbits/sec    1    219 KBytes
[  4]   7.00-8.00   sec  3.24 MBytes  27.2 Mbits/sec    0    233 KBytes
[  4]   8.00-9.00   sec  3.64 MBytes  30.5 Mbits/sec    0    240 KBytes
[  4]   9.00-10.00  sec  3.23 MBytes  27.1 Mbits/sec    0    240 KBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Retr
[  4]   0.00-10.00  sec  32.3 MBytes  27.1 Mbits/sec    5             sender
[  4]   0.00-10.00  sec  31.5 MBytes  26.4 Mbits/sec                  receiver

iperf Done.

--- 10.10.10.1 ping statistics ---
26 packets transmitted, 16 received, 0% packet loss, time 15027ms
rtt min/avg/max = 40.792/53.921/68.759 ms
```

Fig. 19. LTE client channel bandwidth test report example (Iperf and ping tools); (rtt - round trip time);

The results obtained for testing each network access technology are compared in Table 2. We mention that the tests were executed individually and independent of other measurements. As expected, best performance is experienced over a fixed line, followed by Wi-Fi and LTE.

We identify the NFV server as the most relevant throughput limitation point of the setup (especially for connections over the fixed line). The limitations are the result of several internal network virtualization layers of the OpenStack Cloud stock distribution. Nonetheless, bandwidth optimizations can be achieved with cloud distribution tuning or hardware acceleration mechanisms.

V. Conclusions

The dNG-POP architecture, based on the characteristics of distributed data centers, is devised to leverage the advantages of SDN and NFV concepts. In particular, the UAG supports dedicated control and user plane VNFs related to access networks/technologies (e.g., vEPC) and common VNFs applicable to any traffic flow regardless of the access networks (e.g., uDPM and uAUT). Even though a series of benefits result from adopting the presented architecture like reduced footprint, rent, cooling and power consumption, etc. further work is required to automate network resource allocation by integrating the setup in an SDN framework.

FMC is seen as one of the key strategies for deploying future 5G networks aiming at satisfying, in a cost-efficient way, the stringent requirements imposed by advanced services. Within the FMC concept, the deployment of a common and unified functional entity, referred to as UAG, allows seamlessly terminating at the IP layer fixed, mobile and Wi-Fi user traffic flows. By doing so, the network

Table 2. Client channel network performance test results

Network access point	Throughput (Mb/s)		Latency - round trip (ms)		
	Uplink	Down link	Min.	Avg.	Max.
LTE 1	43.5	45.4	16.94	18.01	21.80
LTE 2	26.4	55.1	40.79	53.92	68.76
Wi-Fi	63.5	72.1	1.72	2.382	3.19
Fixed line	676	781	0	0	1

is simplified (not need to maintain and operate independent core networks for each access technology) which does enhance OpEx and CapEx critical for next generation networks.

Our implementation of the UAG concept, with all the required VNFs, has been successfully validated through experiments presented, targeting both the control and data planes. Fixed, mobile and Wi-Fi access users create their sessions demonstrating the FMC capability of the UAG. To this end, a common authentication process (uAUT) for any service is provided. Data path management and content caching capabilities are validated through various use cases that have proven an increase in QoS offered and in user mobility. Last but not least, the UAG provides an attractive platform for exploiting the network sharing concept between different network operators.

Acknowledgment

This work has been funded by the EU ITN Marie Curie project ABACUS – 608363 and FP7 IP COMBO project – 317762.

References

[1] B. Andrus, V. Mehmeri, A. Autenrieth, J.J. V. Olmos, and I. T. Monroy "Evaluation of SDN Enabled Data Center Networks Based on High Radix Topologies," Proc. of The 12th International Conference on Systems and Networks Communications (ICSNC), 2017.

[2] I. Narayanan, A. Kansal, A. Sivasubramaniam "Right-Sizing Geo-distributed Data Centers for Availability and Latency," 37th International Conference on Distributed Computing Systems (ICDCS), 2017

[3] China Mobile Research Institute, "C-RAN: The Road Towards Green RAN", White Paper, 2011

[4] EU funded ICT project COMBO found at: http://www.ict-combo.eu/

[5] A. Gupta, R. K. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies", IEEE Access, vol. 3, pp. 1206-1232, August 2015.

[6] S. Pachnicke, et al., "The Frontiers of Optical Access Networks", International Conference on Optical Network Design and Modeling (ONDM), Pisa, Italy, May 2015.

[7] ICT-COMBO alliance White Paper "A Universal Access Gateway for Fixed and Mobile Integration", April 2016.

[8] S. Pachnicke, B. Andrus, "Impact of Fixed-Mobile Convergence", International Conference on Optical Network Design and Modeling (ONDM), Invited, Cartagena, Spain, May 2015.

[9] S. Gosselin, et al., "Fixed and Mobile Convergence: Which Role for Optical Networks?" Journal Of Optical Communications and Networking (JOCN), vol. **7**, no. 11, 2015.

[10] Z. Tayq, et al, "Fronthaul Performance Demonstration in a WDM-PON-Based Convergent Network", submitted to European Conference on Networks and Communications (EuCNC), Athens, Greece, June 2016.

[11] Common Public Radio Interface (CPRI) - Interface Specification (V.6.0), August 2013.

[12] N. McKeown et al: "OpenFlow: enabling innovation in campus networks", ACM SIGCOMM: Computer Communication, Volume 38 Issue 2, pp. 69-74, 2008.

[13] M. Bjorklund, et al, " Network Configuration Protocol (NETCONF)", RFC 6241, June 2011, found at: https://tools.ietf.org/html/rfc6241

[14] 3GPP Technical Specification 23.335, User Data Convergence (UDC), Stage 2, Release 9, June 2010

[15] T. Mamouni, et al., "Universal AAA for hybrid accesses", EuCNC, Paris, 2015.

[16] A. Ford, et al., "TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6824, Jan. 2013.

[17] Erik Nygren, et al., "The Akamai network: a platform for high-performance internet applications". SIGOPS Operating Systems Review Vol. 44, No. 3, pp. 2-19, August 2010 (DOI: 10.1145/1842733.1842736).

[18] Claudio Imbrenda, et al., "Analyzing cacheability in the access network with HACkSAw.", International Conference on Information-Centric Networking (ICN '14), Paris, France September 2014 (DOI: 10.1145/2660129.2660135).

[19] Salah-Eddine Elayoubi and James Roberts. "Performance and Cost Effectiveness of Caching in Mobile Access Networks", ICN '15, San Francisco, USA, September 2015 (DOI: 10.1145/2810156.2810168).

[20] Giyoung Nam and Kyoung Soo Park. "Analyzing the effectiveness of content delivery network interconnection of 3G cellular traffic". International Conference on Future Internet Technologies (CFI '14), Tokyo, Japan, June 2014 (DOI: 10.1145/2619287.2619297).

[21] EU ICT-COMBO project deliverable 6.3, "Report describing results of operator testing, capturing lessons learned and recommendations", 27 July 2016, http://www.ict-combo.eu/data/uploads/deliverables/combo_d6.3_wp6_2016-07-22_adva_v1.0.pdf

[22] Z. Tayq, et al., "Fronthaul Performance Demonstration in a WDM-PON-Based Convergent Network", European Conference on Networks and Communications (EuCNC), Athens, Greece, June 2016.

Paper C: Live Migration Downtime Analysis of a VNF Guest for a Proposed Optical FMC Network Architecture

B. Andrus, A. Autenrieth, S. Pachnicke,J.J. V. Olmos, and I. T. Monroy "Live Migration Downtime Analysis of a VNF Guest for a Proposed Optical FMC Network Architecture," in *Proc. of 17th ITG-Fachtagung Photonische Netze Conference*, Leipzig, Germany, 2016.

Live Migration Downtime Analysis of a VNF Guest for a Proposed Optical FMC Network Architecture

B. Andrus [1, 2], A. Autenrieth[1], S. Pachnicke[3], J.J. Vegas Olmos[2], and I.Tafur Monroy[2]
[1] ADVA Optical Networking SE, Fraunhoferstr. 9a, 82152 Martinsried/Munich, Germany, bandrus@advaoptical.com
[2] Technical University of Denmark, Ørsteds Plads 343, Kgs. Lyngby, 2800, Denmark,
[3] Kiel University (CAU), Kaiserstr. 2, 24143 Kiel, Germany.

Kurzfassung

Fixed Mobile Convergence (FMC) bedeutet die Verwendung einer gemeinsamen optischen Zugangsnetz-Netzinfrastruktur, die in der Lage ist, festen und mobilen Datenverkehr einschließlich Wi-Fi, Mobilfunk- und Ethernet zu übertragen. Network Function Virtualization (NFV) ist eine wesentliche Voraussetzung für FMC. Die (Live-) Migration von virtuellen Funktionen ist ein weiteres Schlüssel-Feature und bietet bessere Lastverteilung, erhöhte Energieeffizienz, Anwendungselastizität sowie andere wesentliche Vorteile. In diesem Beitrag wird die Auswertung der Migration einer virtuellen Netzfunktion über eine FMC-Infrastruktur präsentiert. Unsere Ergebnisse zeigen, dass die Durchführung einer Live-Migration über eine dedizierte Verbindung keine Ausfallzeiten hervorruft und eine maximal zulässige Verzögerung einhalten kann. In einem weiteren Szenario, in dem die laufenden Verbindungen auf einen anderen optischen Weg umgeleitet wurden, konnte der erfolgreiche Abschluss der Migration mit einer geringfügigen Zunahme der Verzögerung um nur 2,4 Sekunden (22% höher als die Benchmark) und einer Ausfallzeit von nur 2 Sekunden gezeigt werden.

Abstract

Fixed Mobile Convergence (FMC) implies use of a shared optical fronthaul network infrastructure able to carry transparently both fixed and mobile traffic including Wi-Fi, Mobile and fixed Ethernet. Network Function Virtualization (NFV) is a main enabler for FMC using a shared infrastructure for fixed and mobile gateways. Live migration, a virtualization key-feature, offers load-balancing, increased energy efficiency, application elasticity and other worthy advantages. This paper presents the evaluation of migrating a VNF over an FMC infrastructure. Our results show that, performing a live migration over a dedicated connection yielded zero downtime and met a benchmark delay. The following scenario, where the ongoing connection is re-routed on a different optical path, shows the successful completion of the migration with an increase in delay of 2.4 seconds (22% higher than the benchmark) and only 2.1 seconds downtime.

1 Introduction

Mobile traffic is expected to increase close to eight fold between 2015 and 2020 on a global scale. Combined with a prediction of Voice Over LTE (VoLTE) surpassing Voice Over IP (VoIP) by 2019 we see a clear demand for a unified service experience.[1] Fixed mobile convergence (FMC) is an upcoming proposed architecture in communications and service provisioning. FMC is regarded as a transition point in the telecom industry aiming at dismissing the borders between fixed and mobile networks. From the point of view of the subscribers, such a convergence translates in a simplified connectivity. Accessing data, video or voice services relieves users of any concern for how services are delivered. In addition, any uncertainty regarding inaccurate service charges can also be eliminated as there is only a unique service provider for the subscriber.

The structural architecture of FMC (infrastructure sharing) will have a direct impact on the evolution of optical networks. Due to passive optical network (PON) architecture, large scale residential deployment of fiber to the home can make use of shared fiber infrastructure. Even though the situation is different when it comes to mobile networks, the role out of a new passive optical access network infrastructure will be required in the future and it will make sense to share this between both mobile and fixed architecture. A possible solution for this pooling and integration can be found in the maturation of coarse wavelength division multiplexing (CWDM) developing a bidirectional communication feature on a single fiber.[2]

Extending on the FMC definition, convergence can be split into three areas: device, service and network convergence. Device convergence means the integration of diverse functionalities into a single device like authentication, content caching, billing services etc. The key for achieving device convergence lies in Network Function Virtualization (NFV) technologies. NFV [3] is the paradigm that focuses on decoupling network functions that are currently tightly embedded into specialized devices with the scope of running them exclusively in software. This concept facilitates the development of a shared infrastructure for fixed and mobile gateways. NFV employs commercial off-the-shelf servers for hosting basic functional modules which can be combined for each service in a flexible way. Network virtualization is part of a broader scope, along with storage and processing resources contributing to a fully virtualized infrastructure for ultimate efficiency and flexibility. Other advantages stemming from virtualization in the networking domain range from better scalability and adaptability to user demands, lower capital investment costs because of comoditized hardware, simplified operations to granular and custom security.

Network convergence complements device convergence in order to establish connectivity to functions and services using the most suitable technology in a specific location or moment in time. An efficient network convergence cannot be done without a common intelligence, having a view over the entire network i.e. a centralized network controller. In addition, the ability of end-to-end path provisioning, flexibly and on demand for each service is required. Such requirements can be identified within the characteristics of Software Defined Networking (SDN) technology. SDN can be reduced to a programmable interface between administrators and network devices decoupling the control plane from the data plane. A series of advantages make SDN a suitable adopted technique in the FMC architecture like: unified view and control of heterogeneous devices, service provisioning, increased security options and granularity etc.

Service convergence refers to the ability of delivering services to the user in a transparent and seamless fashion over any available network. This requires a flexible deployment of network functions across the network according to user mobility and traffic shifts. The ability to move services and functions through the network is a key feature of virtualization. Migration of a virtual machine (VM) is regarded in computer science as the act of moving a VM from one host to another. Live migration assumes an additional constraint on this feature; the transfer must be executed without considerable service interruption and affecting availability. This feature, which is essential to the cloud environment, can easily be adopted in the NFV application domain. Benefits brought by such an adoption can be translated into traffic load balancing, failure protection, point of presence shift, application elasticity and more.

The following paper focuses on presenting VNF migrating use cases, requirements and concerns introduced by migration as well as analyzes the performance of migrating a VNF between two NFV servers over an FMC compliant network.

2 VNF live migration considerations

As mentioned in the previous section, live migration brings a series of advantages to the NFV table:

- Load balancing user traffic across the access network can be achieved by transferring network functions between different processing nodes.
- restoring a VNF with an easy to create and replace copy leads to hardware failure protection and service continuity.
- scaling the VNF graph enhances application elasticity and shifting the point of presence to a local premises allows VNF debugging

In a converged access network, LTE and potentially similar 5G procedures play an important role in the order of how functions must process the user traffic [4]. Network services like attaching, tracking, paging, data calling, content filtering not only require specific processing functions but also a strict order in which they are processed. The processing path for a network service is defined by a forwarding graph or a VNF graph. For example, when setting up an LTE bearer, a tunnel must be created between the user equipment (UE) and the service gateway (S-GW). The one responsible for creating the tunnel is the evolved node B (eNB) to which the UE is connected. User traffic is transported over IP and the tunnel must have the IP destination of the S-GW virtual function. In order to achieve flexible and scalable applications, enabled by virtual function mobility, the forwarding graph needs to be updated in accordance with the migration.

A VNF migration not only moves the control plane but also the data plane and the service network into a cloud based network. The migration must present stable service continuity, availability, resiliency in both control plane and data plane[5]. A possible method can involve maintaining all the VMs in a forwarding chain in a single startup procedure. By employing this method, the whole graph can be migrated manually or according to known network tidal effects e.g. daily traffic shift from residential to business areas and vice versa.

Any change in the hardware can influence the live migration configuration. Therefore, executing the migration without service interruptions implies details about the workings of the VM hypervisor and the underlying hardware as well. A live migration requires the transfer of the VM with all the components as well as its running state. Regularly the state is stored in the hypervisor and the host kernel as well. All states need to be transferred with minimum downtime in order to maintain the live characteristic. Two operations need to be taken into consideration when live migrating: the virtual IO state and the memory data transfer.

2.1 State of the art

There are many research projects focusing on NFV and the possible benefits virtualizatoin may add [6], [7]. Advanced research in the virtualization field has mostly focused on live migration performance in cloud computing environments [8]. However, an in depth analyses on the implications of VNF live migration, especially with regard to a FMC network architecture, has not been pursued.

3 Scenario

In the following section we consider an FMC aggregation network with a proposed universal access gateway (UAG). The UAG plays the role of a next generation Point-of-Presence (NG-POP) supporting mobile and fixed network functions. Examples of network functions to be hosted on the UAG and desirable for FMC are:

- universal authentication authorization accounting (uAAA) for clients roaming seamlessly from a mobile network to a wireless one and vice versa
- content caching that would facilitate client access to streaming and other web services while offloading the core network of additional traffic
- mobile EPC bearing the core network of the LTE system (MME, S-GW, PDN-GW).

The UAG presents the ability to route the user traffic, according to the source, towards a pool of base band processing units (BBU hotel) or towards the packet core[9].

Extending on this setup, we consider two scenarios that

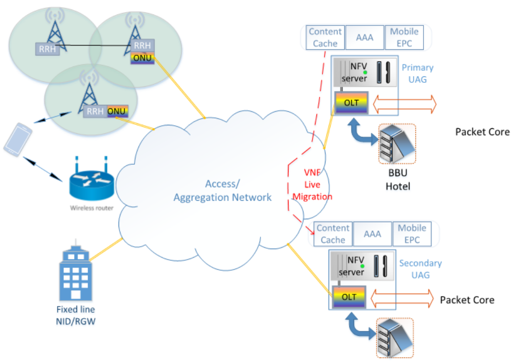


Figure 1 VNF live migration as a backup scenario in an FMC access network

would focus on highlighting the versatility of executing a VNF live migration. In a first use case, we consider the access network having two entry points (UAGs) towards the packet core. One primary UAG bearing the main responsibility for the user traffic and a secondary one that plays the role of a backup PoP. The secondary UAG can either process traffic from temporary events like sports stadiums, conference/business centers or be powered down in the periods of low utilization. The migration can occur as part of a load balancing architecture or redundant connection to the core. A graphical representation of this scenario can be seen in fig. 1. As a backup implementation, in the event of a failure detected on the primary UAG, a fail-safe action of live migrating the whole forwarding graph to the backup UAG can be taken. The advantage of live migration would translate into minimum downtime of the user traffic, which needs to be rerouted towards the secondary UAG.

As a second scenario, we consider an access network connecting to the packet core through a UAG. On the customer premises the service provider can deploy an NFV server hosting default services like firewalling, encryption, scrubbing etc. In order to optimize network utilization, content caching, performance monitoring or path selection mechanisms can be migrated closer to the client devices.

Minimizing the functional downtime during the transfer is

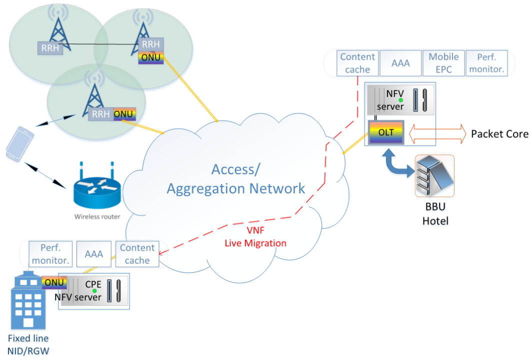


Figure 2 VNF live migration as part of traffic optimization

a high priority. A minimum latency and high bandwidth must be secured for the migration event. A logical solution is to provision a temporary path over the low latency access network. In order to achieve this, an SDN controller having a centralized view and control over the network, can create/tear-down of a path through the network from source to destination. By exposing a northbound API, a VNF manager as part of a service orchestrator monitoring the service statistics, can trigger the migration on the server side and the path provisioning through the SDN controller. The functional overview of such a scenario can be seen in fig. 3

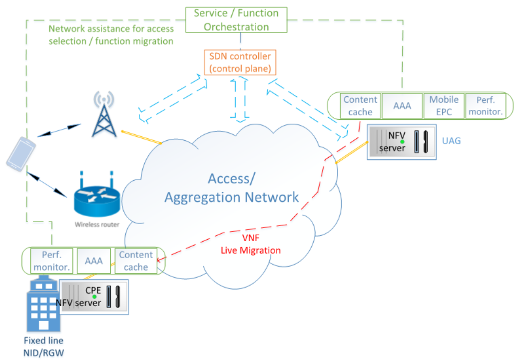


Figure 3 Functional convergence

4 Test setup and measurements

4.1 Experimental setup

In order to test the implications of executing a VNF live migration over an FMC capable optical network we deployed the setup in fig. 4. The low latency optical switch has a custom SDN implementation based on Netconf and YANG modeling. OpenDaylight, the SDN controller employed for the setup uses a Netconf connector as a client in order to connect to the low latency switch. By extracting the capabilities and configuration it is able to control the device. The northbound Rest interface, provides a means of communication with a higher function orchestration script. The orchestration script is responsible for triggering the migration on the remote NFV servers and provisioning the path through the low latency switch. The two remote NFV servers are hosting multiple virtual machines representing network functions. The underlying virtual machine hypervisor handling the migration inner workings is KVM. In order to show the impact of rerouting an optical path, two optical network interfaces were employed on the right hand server from fig. 4.

A 1Gbps Ethernet link was configured across the network

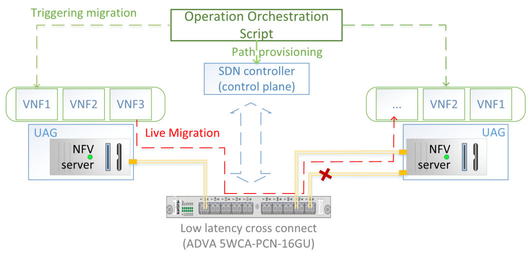


Figure 4 VNF live migration experimental setup

from one server to the other. A control plane connection from the SDN controller to the low latency switch and the hypervisors was established over the management network. IP addresses for all virtual machines and NFV servers were allocated in the same subnet. The reason for this was to simplify the migration operation and reducing additional downtime imposed by a VNF address change.

4.2 Measurements

For establishing a benchmark on the migration performance we started and migrated a VM with a size of 1GB. Optical fibers in the setup had a length of 3m. On the path, no additional traffic was inserted. For quantifying the performance we used two relevant metrics for such an operation. Measuring the migration delay represents an important standard showing mainly the influence of the network and the bandwidth. Capturing the difference between the start and the end of the operation was achieved with the help of the hypervisor tools. Another relevant metric is the VM downtime, the period in which the VM is not responsive during the migration. According to documentation[10], a downtime testing procedure cane be done by executing a high frequency series of pings. The interval between consecutive pings towards the target machine was selected at 50 ms. The downtime is recovered by multiplying the number of lost (non responsive) pings with the selected interval to be of 50 ms. The measurement accuracy resolution for the downtime is 50 ms.

We can observe in 1 that our benchmark migration over a fixed optical link of 1Gbps took an average of 10.78 s. In the following scenario in which we simulated a path switch on the low latency cross connect level, an average additional operational delay of 2.4 seconds was measured.

Migration Delay	**Min.**	**Aver.**	**Max.**
Fixed path	10.159	10.784	11.175
Switched path	12.507	13.206	13.661

Table 1 Migration delay measurements

Tuning the migration, results in a zero downtime for

Migration Downtime (s)	**Min.**	**Aver.**	**Max.**
Fixed path	0	0	0
Switched path	1.8	2.1	2.45

Table 2 Migration downtime measurements

our benchmark testing. In table 2 we can observe that changing the optical path incurs an average of 42 lost packets which translates to 2.1 s downtime. The lowest count of lost packets was 36 out of the total implying a 1.8 s downtime, and the highest was measured to be 2.45 s.

5 Conclusions

This work focused on identifying and examining the issues and considerations regarding VNF live migration with regard to an FMC architecture. Moreover, complementing this section, we performed a series of measurements equivalent to possible VNF migration scenarios. From the results we concluded that an FMC capable optical network allowed the migration of a VNF between two NFV servers with zero downtime. In the event of the path being switched at the optical level, the migration performance is affected by an additional 2.42 s delay representing a 22.45% increase in migration execution. However, an average of only 2.1 s downtime is also introduced by the switch which does not affect service level agreements defined by most service providers. If zero downtime is required during the migration, enabling SDN control over the network elements achieves the goal. Through this setup, a temporary fixed path can be provisioned for the duration of the transfer by a higher service orchestrator.

5.1 Future work

The next logical step for extending this research is the analysis of the effects of migrating a VNF on the forwarding chain. As mentioned above, migrating all the VNFs in the graph could provide a solution, however, compromises on flexibility would need to be made. Another aspect could focus on enhancing the software performance of the migration by testing light weight containers for hosting network functions.

6 Acknowledgment

The research leading to these results has received funding from the European Union Marie Curie ABACUS and COMBO projects.

7 Bibliography

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper

[2] Gosselin, S. et. al.: Fixed and Mobile Convergence: Which Role for Optical Networks?, Optical Communications and Networking, IEEE/OSA Journal of (Volume:7 , Issue: 11), p. 1075-1083, Nov. 2015

[3] European Telecommunications Standards Institute, ETSI - NFV, 2016, `http://www.etsi.org/technologies-clusters/technologies/nfv`, accessed 14-February-2016

[4] 3GPP, Technical Specifications 32.426, V10.3.0, March 2011, `ftp://www.3gpp.org/TSG_SA/WG5_TM/TSGS5_76/_specs_for_checking/32426-a30.doc`, accessed 14-February-2016

[5] IRTF, Elasticity VNF, Oct. 2014, Internet draft, `https://tools.ietf.org/html/draft-zu-nfvrg-elasticity-vnf-00#section-4.2.2`, accessed 14 February 2016

[6] Myung-Ki Shin et al, Verification for NFV-enabled network services, International Conference on Information and Communication Technology Convergence (ICTC), 2015

[7] R. Bonafiglia et al, Assessing the Performance of Virtualization Technologies for NFV: A Preliminary Benchmarking, Fourth European Workshop on Software Defined Networks (EWSDN), 2015, Pages: 67 - 72

[8] Syed A. R. Shah et al, A performance analysis of precopy, postcopy and hybrid live VM migration algorithms in scientific cloud computing environment, International Conference on High Performance Computing & Simulation (HPCS), 2015

[9] A. Magee et al, A Modern Approach to Fixed Mobile Convergence on the Universal Access Gateway, 16th IEEE International Conference on High Performance Switching and Routing, Budapest, 2015

[10] Felix Salfner, Peter Troger and Matthias Richly , Dependable Estimation of Downtime for Virtual Machine Live Migration, International Journal on Advances in Systems and Measurements, vol 5 no 1 & 2, 2012

Paper D: Performance Evaluation of Two Highly Interconnected Data Center Networks

B. Andrus, O. M. Poncea, J.J. V. Olmos, and I. T. Monroy "Performance Evaluation of Two Highly Interconnected Data Center Networks," in *Proc. of 17th International Conference on Transparent Optical Networks (ICTON)*, 2015.

PERFORMANCE EVALUATION OF TWO HIGHLY INTERCONNECTED DATA CENTER NETWORKS

Bogdan Andrus(1), Ovidiu Mihai Poncea(2), J.J. Vegas Olmos(1), I. Tafur Monroy(1)
(1) Technical University of Denmark, Ørsteds Plads 343, Kgs. Lyngby, 2800, Denmark
(2) Faculty of Automatic Control and Computers, University POLITEHNICA of Bucharest, Romania
(1) Tel: +45 45 255 170, e-mail: bogan@fotonik.dtu.dk

ABSTRACT
In this paper we present the analysis of highly interconnected topologies like hypercube and torus and how they can be implemented in data centers in order to cope with the rapid increase and demands for performance of the internal traffic. By replicating the topologies and subjecting them to uniformly distributed traffic routed by shortest path algorithms, we are able to extract relevant statistics related to average throughput, latency and loss rate. A decrease in throughput per connection of only about 5% for the hypercube compared to 16% for the 3D torus was measured when the size of the network was increased by a factor of 32. The performance measurements are supported by abstract metrics that also give a cost and complexity indication in choosing the right topology for the required application.
Keywords: hypercube, torus, interconnection networks, NS3, Data Center topologies, performance evaluation.

1. INTRODUCTION

According to the projections of Cisco Global Cloud Index 2014, the Data Center Traffic is expected to grow by roughly a factor of three times by 2018 to an amount of 8.6ZB; out of this 75% fits into the category of east-west traffic inside the Data Center [1]. This prediction puts more and more pressure on the internal network that needs to support a growing amount of influx packets travelling between servers in dynamically changing traffic patterns. An interconnection topology plays an important role for Data Center applications and efficient on-chip networks alike [2]. For this reason, topologies initially designed for parallel computing and on-chip networks, that present higher interconnectivity between nodes, have started to be increasingly adopted in Data Center networks [3]. A couple of good examples that fit this need, and are well known for their performance, are hypercube and torus which will be comparatively analysed in detail in this paper.

Another crucial aspect that is determined by the topological structure of a network is reliability. The number of redundant paths, but most importantly their placement in the overall layout increases the networks ability to experience local failures without major impact on operations. In addition, it regulates the latency by introducing a number of devices (nodes) along the path of a message that have to be traversed in order for the message to reach its destination. Regardless of the scope, the desired characteristics of a network revolve around high accepted throughput per input [4], low latency, high reliability and scalability and all of this at the lowest cost possible. However quantifying true values for these criteria is often hard and usually is based on probabilistic models.

The early stage design of a network cannot be based on real system experiments when it comes to large dimensional models. However, taking advantage of a highly controllable environment, such as the one found in a network simulator, is a good compromise. NS3 is used for modelling the most important components of a system which can undergo a series of simulation types that lead to trustworthy and reproducible results. This replication of the real hardware structure mimics the observable behaviour and gives a close approximation of the desired parameters.

2. THEORETICAL CONSIDERATIONS IN NETWORK MODELLING

A key component in establishing the effectiveness of a Data Center architecture is the network. The impact of a topology on the global network ratio of performance vs. cost is remarkable. Traversing nodes and links incurs energy, and since the number of hops for the various paths is affected by the interconnection implementation an important role in the energy consumption can be clearly identified. In this section we will evaluate how, and to what extent, these statements apply for the torus and hypercube topologies.

A hypercube implementation is an n-dimensional generalization of a 3-cube also called n-cube and comprises of 2n nodes (switches). It is known to have a high connectivity and small diameter however, not very easy to scale due to complexity. A 16 nodes hypercube is displayed in *Figure 1*. A three-dimensional torus can be visualized as a three-dimensional mesh interconnect in the shape of a rectangular prism with all the nodes on each face having an additional connection to the corresponding nodes on the opposite face, as portrayed in *Figure 2*. Torus networks can be found in top performing supercomputers due to their low cost and reduced diameter compared to mesh.

A few of the most relevant mathematical parameters by which topologies can be characterized and compared in a preliminary Data Center network design stage have been calculated in similar research [5]. With the use of the following properties we can build a sanity-check standard for clarifying certain expectations, in relation to scalability for each network but also in relation to each other.

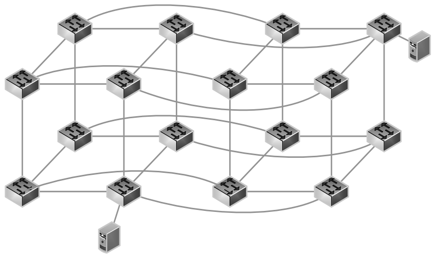


Figure 1. 16-Node Hypercube.

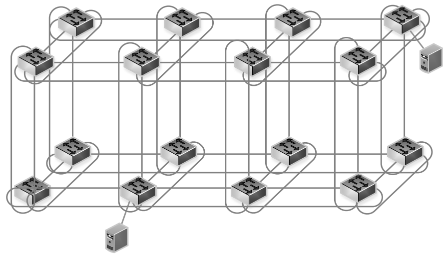


Figure 2. 16-Node Torus.

Bisection bandwidth, defined as the total bandwidth across a cut through the middle of the network, gives the link density and bandwidth that can be achieved by a certain implementation strategy. We can see in *Figure 3*, that even though the order of growth is similar in both implementations, the bisection reaches larger values for hypercube. This asset of superior bandwidth comes with a setback related to cost and complexity of wiring. However, this difference is almost inexistent for 64 nodes and below.

The *diameter* of the network, the longest path between any two nodes calculated on the shortest path tree, is arguably a sign of packet latency. Another similar guideline is *average distance* in the network that also supports latency. A similar tendency is observed for both parameters in *Figure 4* and *Figure 5* and because a network with a smaller diameter is more desirable, the hypercube displays a favourable growth as compared to the torus.

Node order represents the number of interconnection links at each node and relates to network throughput however, implementing a system with a high node order implies an increased execution complexity cost. A relevant observation, by depicting the progression of the node order in *Figure 6*, is that the torus network can scale up and maintain a constant node degree if the dimension of the implementation strategy does not change. Again, the threshold of 64 nodes is where the balance shifts between topologies and the hypercube begins to surpass the torus with a logarithmic rate).

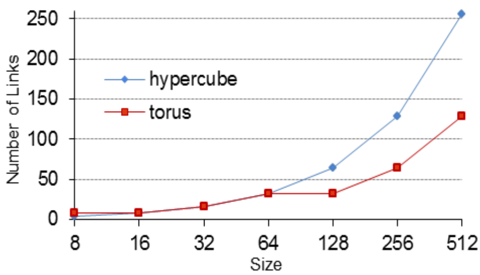


Figure 3. Bisection Bandwidth.

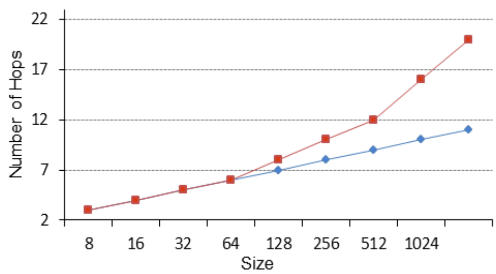


Figure 5. Diameter.

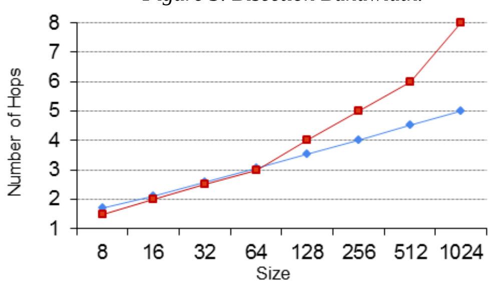


Figure 4. Average Distance.

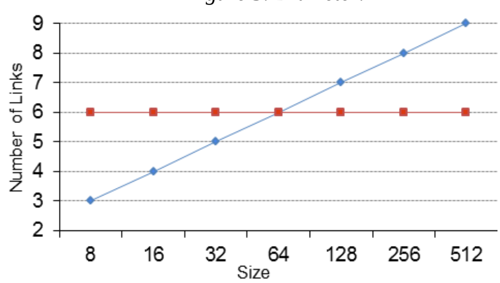


Figure 6. Node Order.

3. NETWORK PERFORMACE SIMULATION

Through the following simulations, we try to provide a holistic interpretation of the highly interconnectable networks examined. Therefore, when building our simulation model, we are considering topology characteristics, communication pattern and the amount of data injected in the network as being the most relevant components for our two scenarios. The implementation and use of these elements is described for an overall understanding of the work and the results.

After determining the topological characteristics of the selected networks, the routing algorithm is selected with the purpose of making decisions for every network node concerning the best paths for the data. Routing algorithms have been studied and developed in various research projects [6],[7]. As far as path computation is concerned, a minimal routing algorithm that computes the shortest paths between any source and destination was used in our simulations. Since all links are presumed homogenous, having the same capacity and delay (length of cables only implies a negligible latency compared to the one caused by processing and congestion), a process selecting only the paths that require the smallest number of hops between any two nodes has been selected. However, link contention may still occur influencing the measurement of connection throughput, latency and loss rate which is analysed in the simulation results.

The communication pattern is another component that has a great influence on performance. Selecting and implementing an appropriate traffic pattern is a concern that has to be tackled. For the following scenarios, we will adopt a uniformly distributed random traffic model that is widely accepted [8] and is unbiased towards various topologies. This is done by a traffic generation module programmed to distribute packets throughout the entire network, assigning at random roles of clients or servers to each host and then selecting arbitrary destinations for the packets travelling between them at various points in time.

Another important aspect that has to be taken into consideration is the amount of data that will be injected in the interconnection network by each node sourced in the local segment. From documentation [9], we know that in the on-chip networks domain the term injection rate is used to measure the rate that the packets are inserted into the network. The issue in this case, is that above a certain threshold the network begins to saturate, the ratio between throughput and injection rate starts to drop and the efficiency decreases. The saturation limit for hypercube and torus is 60% and 40% respectively [10]. Adopting this characterization for an IP network, we would consider the input traffic level to be at 30% the maximum link capacity, well below the saturation limit.

In order to simulate the proposed scenarios we used NS3, an open source discreet event network simulator, based on C++ programing language. NS3 platform is used by network, routing protocol designers and researchers who wish to asses and classify large sets of information regarding performance metrics related to throughput, delay and number of lost packets.

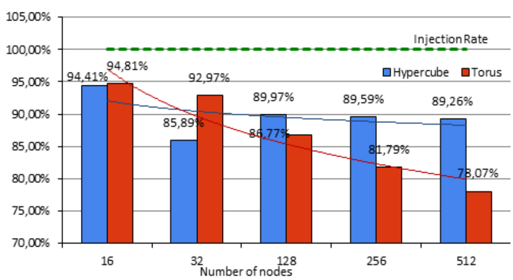


Figure 7. Throughput per Node.

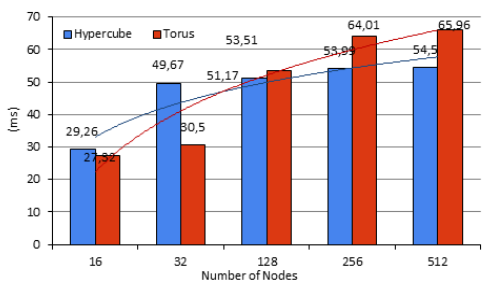


Figure 8. Average Latency.

For data center architects, a more relevant measurement is the evolution of available bandwidth per node as the network scales, compared to the total network bandwidth. Plotting the results of the proposed simulation model, it is easy to observe that, as the number of nodes grows, the resource contention also gets higher and visibly the bandwidth allocation per node decreases – *Figure 7*. This statement is also backed by focusing on the evolution of average latency and lost packets per connection – *Figure 8*, *Figure 9*.

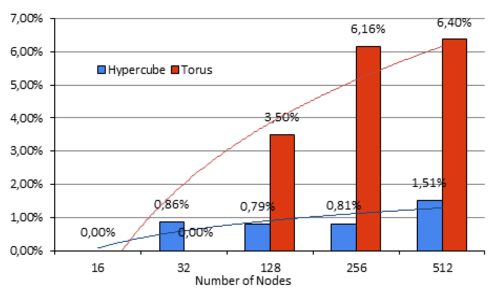


Figure 9. Lost Packets Rate.

4. CONCLUSIONS

Data center network topologies display a rich and diversified design space. Through this paper, we have shown how the performance of the studied interconnections decreases as the network scales up to 512 nodes and the bandwidth requirement of servers for intercommunication remains the same. Even though the ejection rate trend is declining, the hypercube topology presents more consistency with a decrease of only 4,86% compared to 15,9% corresponding with the torus. Same behaviour is demonstrated for latency and number of lost packets.

A comparative infrastructure cost analysis for implementing the packaging strategies reveals that there is a shift between the two, with the torus being more costly than the hypercube for networks below 64 nodes and the hypercube having a faster growth rate beyond this limit. This conclusion is based on the total number of links, bisection width - Figure 1 and node degree - Figure 4.

Hypercube would be the better choice out of the two, when it comes to large networks, where performance-critical Data Center applications and massive parallel data-crunching processes like Big Data, server operational clusters or even biochemical science, physics, astronomy or climate research, where the aggregate machine needs to present high speed interconnections to keep all the components working together at their peak effectiveness. However, just as we predicted in the early analysis stage of the bisection bandwidth, average distance and node order, the tours is a viable choice when the size does not go beyond 64 elements, presenting better throughput levels per node. This can be explained by the higher node degree and slightly smaller average distance for this interval displayed by the torus.

ACKNOWLEDGEMENT

This work has been partially funded by the Marie Curie project ABACUS and Marie Curie FENDOI.

REFFERENCES

[1] *Cisco Global Cloud Index,* 2014.

[2] P. G. C. L. Jie Chen, “Performance Evaluation of Three Network-on-Chip Architectures (Invited),” in *First IEEE International Conference on Communications in China: Communications QoS and Reliability (CQR)*, 2012.

[3] J. K. Dennis Abts, in *High Performance Datacenter Networks,* Morgan & Claypool, 2011.

[4] N. Enright, On-Chip Networks, Morgan and Claypoo, 2009.

[5] W. J. Dally, “PERFORMANCE ANALYSIS OF K-ARY N-CUBE INTERCONNECTION NETWORKS,” *IEEE TRANSACTIONS ON COMPUTERS,* vol. 39, no. 6, pp. 775-785, 1990.

[6] A. Bossard and K. Kaneko, “On hypercube routing and fault tolerance with bit constraint,” *Second International Symposium on Computing and Networking (CANDAR),* pp. 40-49, 2014.

[7] M. A. Khan and A. Q. Ansari, “Quadrant-based XYZ dimension order routing algorithm for 3-D Asymmetric Torus Routing Chip (ATRC),” *International Conference on Emerging Trends in Networks and Computer Communications,* pp. 121-124, 2011.

[8] R. A. Raghunath M.T., "Designing interconnection networks for multi-level packaging," in Supercomputing '93. Proceedings, 1993.

[9] M. A. R. Mostafa Haghi, "Evaluation of Effect of Packet Injection Rate and Routing Algorithm on Network-on-Chip Performance," IJIRSET , vol. 3, no. 2, pp. 9589-9597, 2014.

[10] A. S. Muhammed Mudawwar, "The k-ary n-cube Network and its Dual: a Comparative Study," in *IASTED International Conference on Parallel and Distributed Computing and Systems*, 2001.

Paper E: SDN Data Center Performance Evaluation of Torus and Hypercube Interconnecting Schemes

B. Andrus,V. Mehmeri, J.J. V. Olmos, and I. T. Monroy, S. Spolitis, V. Bobrovs "SDN Data Center Performance Evaluation of Torus and Hypercube Interconnecting Schemes," in *Proc. of Advances in Wireless and Optical Communications (RTUWO)*, 2015.

SDN Data Center Performance Evaluation of Torus and Hypercube Interconnecting Schemes

Bogdan Andrus, J.J. Vegas Olmos, Victor Mehmeri, Idelfonso Tafur Monroy
Department of Photonics Engineering, Technical University of Denmark, Ørsted Plads, Building 343, Kgs. Lyngby, 2800, Denmark
bogan@fotonik.dtu.dk

Sandis Spolitis, Vjaceslavs Bobrovs
Institute of Telecommunications
Riga Technical University
Riga, Latvia
Sandis.spolitis@rtu.lv

***Abstract*— by measuring throughput, delay, loss-rate and jitter, we present how SDN framework yields a 45% performance increase in highly interconnected topologies like torus and hypercube compared to current Layer2 switching technologies, applied to data center architectures.**

Keywords*— *software defined networking, traffic engineering, optical networks, data centers.

I. INTRODUCTION

With the current evolution of global data centers, traffic is predicted to grow by a factor of roughly three times by 2018 reaching a whopping 8.6ZB [1]. Out of this 75% fits into the category east-west traffic, internally. This leads to expectations that more exotic topologies, known for their high server interconnectivity, like hypercube or torus will present more and more interest to data center architects. However, implementing networks that could make use of the full potential of such topologies would require numerous network devices capable of intelligently routing or switching packets. Unfortunately the performance and utilization in most cases may not be worth the complexity, the relatively higher implementation and Operational Expenditure (OPEX) costs that such designs impose.

Software Defined Networking (SDN) technology promises to bring programmability that meets application needs in real time and solutions that greatly reduce the set-up time. Another goal is lowering the costs by taking out the architectural control elements from each network interconnection node and centralizing their functionality into a controller responsible for traffic decisions and intelligence for the entire network.

By designing data center networks on the principles of SDN, the resulting architectures can benefit from several advantages: highly flexible and interconnected networks intelligent enough to take advantage of the multitude of redundant links comprising such topologies; another aspect involves lowering the configuration and management complexity imposed by scaling such intricate interconnections by maintaining a general overview and supervision of the whole network inside a dynamic controller. Minimizing the expenses appointed by the need to employ numerous nodes able to support complex path computation algorithms and replacing them with bare-metal switches that provide fewer features but enable low-cost and flexible alternatives is a key point in this discussion.

II. THEORETICAL CONSIDERATION AND RELATED WORK

Highly elaborate interconnecting schemes like hypercube and torus were initially deployed and evaluated as Networks-on-Chips [2], a communication subsystem for an integrated circuit for their high bandwidth, low latency and low power consumption. For these reasons we can observe the benefits of transitioning such designs into specialized data center architectures that need to fulfil similar objectives.

A hypercube structure can be defined as a graph whose node set consists of the 2n Boolean vectors of dimension n, where two nodes are neighbours if the difference between them is exactly one coordinate [3]. A graphical representation is shown in Fig. 1a. Hypercube can also be regarded as a generalization of a cube with 3 dimensions. The mathematical properties of the hypercube are easily represented proving that the number of neighbours a node has and the diameter (maximum distance between any pair of nodes) are both equal to the dimension n.

A torus interconnect can be visualized as a mesh with all nodes on the edges having an added connection to the corresponding nodes on the opposite edge. A 16-node torus can be seen in Fig. 1b. Due to these wraparound links, the diameter is reduced by almost a half compared to the mesh. The node degree (number of neighbours of a specific node) is constant as the network scales up giving the torus an advantage when it comes to implementation costs. However this may indicate a potentially lower performance level having fewer possible routes than the hypercube.

For a network, valuable information can be inferred by analysing parameters like the ones described above but also the number of ports, designated for switch-to-switch interconnection, has an influence on cost estimation and implementation complexity [4]. Another key characteristic, for both of the topologies, is edge-symmetry in which all nodes have the same degree. Because of symmetry the network can balance traffic across different channels and decrease the possibility of bottlenecks.

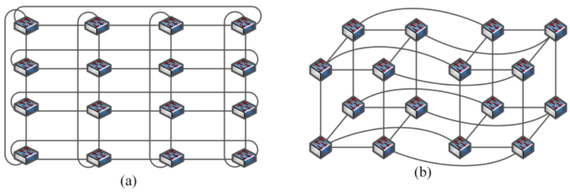


Fig. 1. Torus (a) and Hypercube (b) topologies.

Applying these designs into large data centers utilizing current routing or switching technologies encounters numerous obstacles. A Layer 2 topology that builds upon high density port switches that process packets at line speed implies less configuration and administration overhead, but does not scale well due to the limitations of a flat topology and restrictions to a broadcast domain. A routing implementation can segment the broadcast domains, use existing routing protocols and present better scaling properties however, routers not only process packets slower and are more expensive but also require more administration [5]. In data centers, generally we see a compromise based on a combination of the two but also replacing network elements with expensive multilayer switches is a different option. However, SDN can truly get the best of both scenarios by giving routing capabilities to lower cost, white-box switches and automate administration operations with a topology manager module in the controller. This results in a network that can scale easily in a plug and play fashion.

III. Simulations and Results

In order to have a baseline for comparing the presented topologies we need to establish a set of performance attributes, relevant for data center architects, which can be measured and analyzed. In this regard, we define the ejection rate (throughput) as the data rate in bits-per-second accepted, processed and delivered to the correct destination by the network. Traditionally, the central metric in a network was bandwidth however, network induced latency also has a substantial impact and cloud storage efficiency is a relevant example. We regard latency as the minimum time a network needs in order to transmit a piece of data from source to destination and the average latency results from averaging the delay measured for all the established connections. The packet loss ratio that is obtained by dividing the number of lost packets to the total number of packets sent, is important not only to QoS traffic but also to various traffic control algorithms. Another relevant measurement, for downstream services in particular, is jitter which can be defined as the variation in time between the arriving packets, caused in our scenarios by network congestion.

The characteristics of the traffic pattern play an important role in designing a data center. Due to the fact that data center traces are not publicly available, in similar projects [2], [6] generation of random traffic patterns was selected. In the uniform random traffic trace we adopted, the source and destination nodes are chosen among all nodes with equal probability with an exception that the source address and destination cannot be the same thus avoiding the effort of preventing one node from sending data to itself.

The level of traffic generated by the processing elements that will be injected in the network has to be taken into account in relation to the link bandwidth. Since torus and hypercube belong to the k-ary n-cube larger network class, we have selected a generation rate of 40%, below the class saturation limit corresponding to a size of 256 nodes.

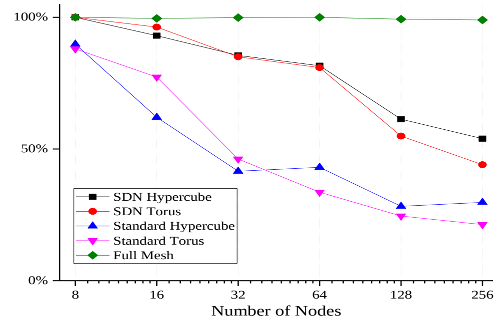


Fig. 2. Throughput per Connection (Ejection Rate).

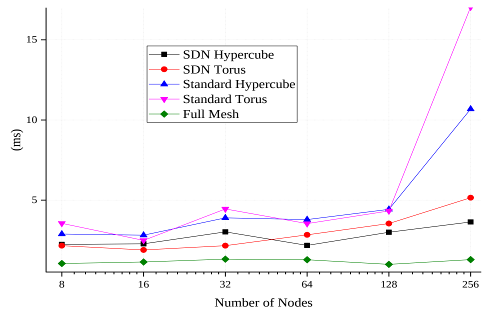


Fig. 3. Delay Performance.

Virtualization infrastructure allows researchers to bridge the gaps between small scale experiments and simulations and real live deployments. For our testing environment, we have utilized Mininet, a virtualization testbed for developing and experimenting with SDN applications. Mininet runs real time, production quality, multilayer virtual switches that can connect to various software controllers and expose realistic network conditions to the applications running on it and even carries traffic on behalf of real users.

To show how an SDN implementation measures against an established Spanning Tree Protocol (STP) enabled network we have evaluated torus and hypercube in both scenarios under the same traffic conditions, using identical virtual switches running on the same platform. To benchmark the fidelity of our testing environment, a full mesh interconnection was assessed and subjected to the same tests.

In order to control the data path distributed across the virtual switches we have employed Floodlight, an enterprise-class SDN Controller. The integrated topology and forwarding module enable the efficient use of the high number of redundant paths for networks like torus and hypercube and calculates connections based on shortest path between each pair of source and destination. However, link contention still occurs, influencing performance as seen in the results.

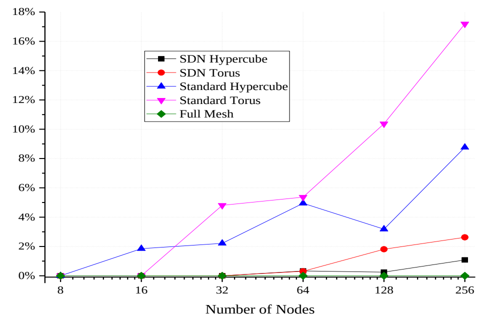


Fig. 4. Loss Rate.

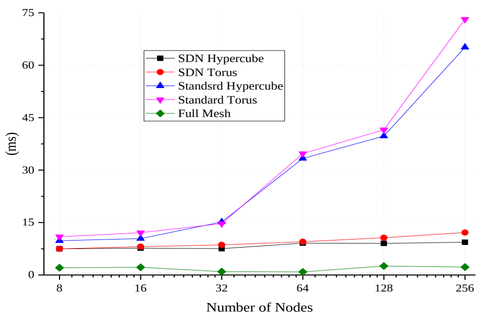


Fig. 5. Jitter Performance.

For data center design, a significant measurement is the bandwidth available per each node as the network expands in number of switches and connections [4]. Normal traffic operation is presumed, as well as a constant injection rate per application running on each server. Since network contention grows, the average throughput for each connection decrease when the networks scale 32 times in size, reaching 256 nodes.

In Fig. 2 we observe the measured throughput is higher by roughly 45% for the SDN test cases with 256 nodes. A 13ms decrease of packet delay, for SDN torus of the same size is noticeable in Fig. 3. Loss rate is also reduced considerably for both topologies with at least 7 percentage points, Fig. 4. Connection stability is improved with SDN technology by a reduction of jitter with 50ms, measurement that applies to networks of 256 switches, Fig. 5.

IV. Conclusions

Implementing an SDN architecture brings a significant performance boost in both torus and hypercube, by taking advantage of the multitude of redundant links. As expected, the growth of the networks and the rise of communication links lead to a decrease in connection performance. However, with the use of SDN technology, networks exhibit better consistency to the scaling effect.

The comparison between the topologies remains similar and applies to both studied frameworks. In this regard, hypercube presents superior performance for larger networks, above 64 nodes. This is also the turning point above which, cost and complexity grow significantly for the hypercube setting the two topologies apart [9]. However, for smaller numbers of switches, performance is similar and torus is a practical choice for deployment.

Acknowledgment

This work has been partially funded by the Marie Curie project ABACUS and Marie Curie FENDOI.

References

[1] Cisco Global Cloud Index, (2014).

[2] J. Chen et al., “Performance Evaluation of Three Network-on-Chip (NoC) Architectures,” ICCC, (2012).

[3] Y. Saad, M. H. Schultz, “Topological Properties of Hypercubes,” IEEE Trans. on Comp., Vol. 37, no. 7, (1988).

[4] B. Andrus, J.J.V. Olmos, I.T. Monroy, “Performance Evaluation of Two Highly Interconnected Data Center Networks,” ICTON, (2015).

[5] K. Jayaswal, Administering Data Centers, Wiley (2005).

[6] M. Al-Fares et al., “Dynamic Flow Scheduling for Data Center Networks,” Proc. NSDI, (2010).

[7] W. J. Dally, “Performance Analysys of k-ary n-cube Interconnection Networks,” IEEE Transactions on Comp., Vol. 39, p. 775-785, (1990).

[8] H. Abachi, A. Walker, “Network Expandability And Cost Analysis Of Torus, Hypercube And Tree MultiProcessor Systems,” SSST, p. 426 - 430, (1996).

[9] M. Haghi et al., “Evaluation of Effect of Packet Injection Rate and Routing Algorithm on Network-on-Chip Performance,” IJIRSET, Vol. 3, no. 2, p. 9589 - 9597, (2014).

Paper F: Evaluation of SDN Enabled Data Center Networks Based on High Radix Topologies

B. Andrus, V. Mehmeri, A. Autenrieth, J.J. V. Olmos, and I. T. Monroy "Evaluation of SDN Enabled Data Center Networks Based on High Radix Topologies," Accepted for publication in *Proc. of The 12th International Conference on Systems and Networks Communications (ICSNC)*, 2017.

Evaluation of SDN Enabled Data Center Networks Based on High Radix Topologies

Bogdan Andrus(1)(4), Victor Mehmeri(1), Achim Autenrieth(4)

(1) Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark.
email: bogan@fotonik.dtu.dk
(4) ADVA Optical Networking SE, Fraunhoferstr. 9a, 2152 Martinsried / München. Germany.

Juan José Vegas Olmos(1)(2), Idelfonso Tafur Monroy(1)(3)

(2) Mellanox Technologies, Denmark.
(3) Department of Electrical Engineering, Technical University of Eindhoven, Netherlands.

***Abstract*— The relevance of interconnects for large future datacenters and supercomputers is expanding as new technologies like Internet of things (IoT), virtual currency mining, High Performance Computing (HPC) and exa-clouds integrate further into data communication systems. The Data Center Network Layer is the workhorse that manages some of the most important business points by connecting the servers between them and delivering high performance to users. Evolution of the networking layer has seen, in addition to improvements of individual link speeds from 10Gb/s to 40Gb/s and even 100Gb/s and beyond, quite important changes in the topological design. Traffic intensive server-centric networks and high performance computing tasks are pushing a shift from the conventional Layer 2 oriented fat tree architectures with multiples tiers towards clos networks and other highly interconnected matrices. Optimal performance and reliability perquisites imposed on the network cannot be fully achieved by solely changing the topological design. A software-centric control of the network enables the use of additional redundant paths not only for increased performance but also reliability concerns. By decoupling the network control from individual devices and centralizing the network intelligence inside a Software Defined Network (SDN) controller, dynamic workloads can easily be accommodated with the deployment of custom modules or applications for traffic management. In this paper, we focus on the innovations for next generation data center networks from a twofold perspective. On the one hand, we evaluate the applicability of new potential interconnection schemes like torus, hypercube, fat tree and jellyfish in regard to identified key metrics such as performance, complexity, cost, scalability and redundancy. Our evaluation comprises of a mathematical interpretation of the graphs with a focus on the abstract metrics (e.g., bisection bandwidth, diameter, port density, granularity etc.) followed by a simulation of the scalable networks in a virtual environment and subject them to various traffic patterns. On the other hand, we introduce an emulated SDN test framework, which decouples the control plane from the interconnection nodes and gives a centralized view of the topology to a controller handling the routing of the internal workflows for the data center. With the use of our SDN enabled testbed we demonstrate and highlight the clear superior performance gain of centralizing the network intelligence inside a software controller, which allows us to apply a custom routing algorithm.**

Keywords- SDN; Data Center topologies; Torus; Hypercube; Jelly Fish; Fat Tree.

I. INTRODUCTION

New technological trends like IoT, virtual currency mining, High Performance Computing (HPC), exa-clouds integrate further and further towards data communication systems making the role of interconnects more important than ever before. The current global evolution of data center traffic is predicted to reach an annual rate of 15.3 zettabytes (ZB) - with a monthly rate of 1.3 ZB - by the end of 2020 [1]. This prediction translates into tripling traffic demands over a period of 5 years spanning from 2015 to 2020 with a compound annual rate of 27%. The distribution of data center related traffic regards the majority of connections established inside the data center with a quota of 77% for server-to-server communication. Major factors influencing such patterns are identified in distributed computing/processing as well as reliable and fast migration of sizable volumes of data across vast domains of physical servers. Such circumstances highlight the importance of the network topology in the process of designing data centers, on account of the fundamental limits (e.g., cost and performance) imposed by the chosen interconnection graph.

Meanwhile, Big Data and Cloud applications, which are subject to exponential growth rates are pressuring Data Center enterprises to drastically improve their infrastructure not just to meet the increased bandwidth demand but also additional QoS perquisites related to certain applications and services. As a consequence, such priorities are placing the boost for network capacity in the top research directions weather they focus on enhancing individual link capacity to 40Gb/s, to 100Gb/s or beyond or by deploying special network topologies and routing structures [2]-[5]. An

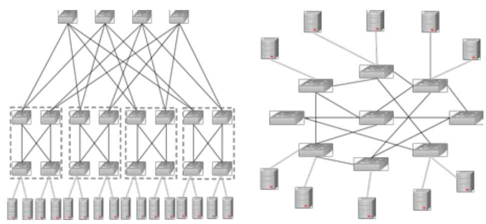


Fig. 1. (a) - Fat Tree (left) and (b) - JellyFish (right) topologies

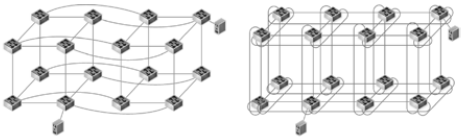


Fig. 2. (a) - Hypercube (left) and (b) - Torus (right) topologies

interconnection matrix is not only an important component in a server-centric Data Center construct but also crucial element for efficient on-chip networks [6]. For this reason, topologies originally adopted in parallel computing platforms and on-chip networks, that present higher interconnectivity between nodes, have gained more and more traction in Data Center network deployments [7].

As identified in [8], one major constraint not tackled by these graphs is the issue of granular scalability or the capability of adding a flexible number of servers or increasing capacity while maintaining structural properties. The number of redundant links, but most importantly their placement in the overall graph raises the network's capability to experience local failures without major impact on operations. The reliability (fault tolerance) of such constructs entails a compromise in terms of additional underutilized spare links or a disproportionate increase of hop count in link failure situations. Furthermore, implications connected to the manual configuration of such networks and potentially complex routing mechanisms not only translate into additional costs related to Capital Expenditure (Capex) but also Operational Expenditure (Opex).

Software Defined Networking (SDN) approach to network management and configuration seeks to bring the flexible programmability needed by real time applications and services, which can considerably reduce the set-up time. One major benefit from adopting an SDN framework relates to lowering Capex and Opex costs. Firstly, low-cost white box switches can be developed by detaching the control plane functionality from all network devices. This step is followed by centralizing their behavior inside a software controller responsible for the management and control operation of the entire network. As such, minimizing the expenses is achieved by replacing the large number of nodes that are capable of supporting complex path computation algorithms with white-box switches that provide fewer features but present a more flexible and reconfigurable alternative. Maintaining a general overview and supervision of every network device inside a dynamic software controller can facilitate overall network management and configuration. Therefore, by automating the manual operations of managing and configuring every network device (also required when scaling such intricate interconnections), Opex oriented costs can also be lowered.

The contributions of this paper can be divided into several sections. In Section 2, we extend our previously presented mathematical analysis [9] of high radix topologies (e.g., Torus, Hypercube) with regard to indications on performance, cost, latency of implementation for new topologies (e.g., Fat Tree, Jellyfish). Section 3 highlights the results and behavior of scaling such topologies in a simulation environment (e.g., [10]) using a random traffic pattern. In Section 4, we present the evaluation testbed for measuring the performance (e.g., network throughput, delay, jitter, and loss rate) of an SDN implementation employing each topology against conventional Spanning Tree Protocol deployment based on our previously published works [11, 12]. Finally, in the Section 5, we present and discuss the results obtained from the simulation and emulation testbeds.

II. BACKGROUND ON NETWORK TOPOLOGIES

A key component in the performance of a Data Center architecture network is the topology. The impact of a topology is not only significant for the global network ratio of performance vs. cost but also for the failure resiliency aspect. Traversing nodes and links incurs energy, and since the number of hops for the various paths is affected by the interconnection implementation, an important role in the energy consumption can also be easily identified.

A hypercube graph, Figure 2(a), is an n-dimensional generalization of a cube also called n-cube and comprises of 2n nodes. One main characteristic is the high connectivity and small diameter however, not very easy to scale due to complexity. A torus topology can be visualized as a three-dimensional mesh in the shape of a rectangular prism with all the nodes on each face having an additional connection to the corresponding nodes on the opposite face, as illustrated in Figure 2(b). Torus based networks are usually employed in top performing supercomputers due to their high radix, relatively low cost and reduced diameter compared to mesh. Another widely deployed Data Center topology is Fat-Tree, Figure 1(a), which is capable of delivering high bisection bandwidth due to its path multiplicity and maintain a low and constant diameter if the number of layers remains constant with scaling. On the other hand, Jellyfish, Figure 1(b), a random graph and multipath based topology, has been proven to be more cost-efficient than a Fat-Tree using identical devices, providing support for 25% more servers running at full capacity [13]. Furthermore, Jellyfish graph provides an attractive solution for a more granular expansion and allows heterogeneity in switch port count, a desired advantage in terms of flexibility.

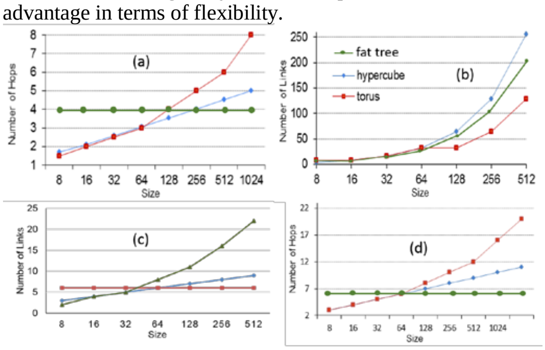


Fig. 3. Average Distance (a), Bisection Bandwidth (b), Node Order (c), Diameter (d)

Some relevant mathematical parameters by which topologies can be characterized and compared in a preliminary network design stage have been identified in

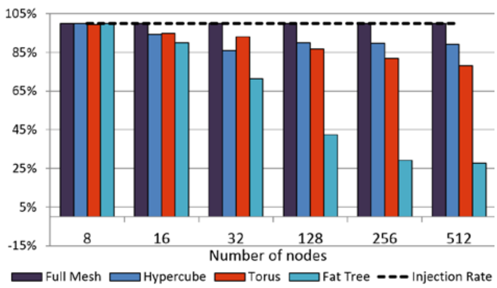


Fig. 4. Average throughput per node

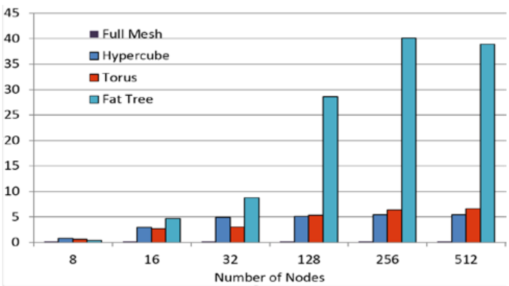


Fig. 5. Average delay per flow

previous research [14]. Such parameters can serve as the building blocks for clarifying certain prospects, in relation to scalability for each network but also in relation to each other.

Bisection bandwidth (Figure 3(b)), the bandwidth sum of all the links across a cut through the middle of the network, gives the link density and bandwidth indications that can be achieved by a certain implementation strategy. Even though the order of growth is similar in torus, hypercube and fat tree, the bisection reaches larger values for hypercube and fat tree. This asset of superior bandwidth comes with a setback related to cost and complexity of wiring on the hypercube side. However, this difference is almost inexistent for 64 nodes and below.

The longest path between any two nodes calculated on the shortest path tree (diameter), is arguably a sign of packet latency, as seen in Figure 3(d). Another similar guideline is average distance in the network, highlighted in Figure 3 (a), which also supports latency in relation to communication patterns. While a similar tendency is observed for both hypercube and torus having a logarithmic and a linear increase, respectively, a 3 layer (e.g., edge, aggregation, core) fat tree maintains a constant diameter if the number of layers is unchanged when expanding.

Node order represents the number of interconnection links (ports) required for each switch and relates to network throughput, however, implementing a system with a high node order implies an increased execution complexity cost. Like in the previous cases, for graphs up to 64 nodes the

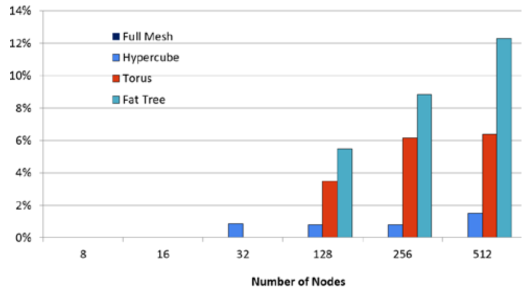


Fig. 6. Average loss rate per flow

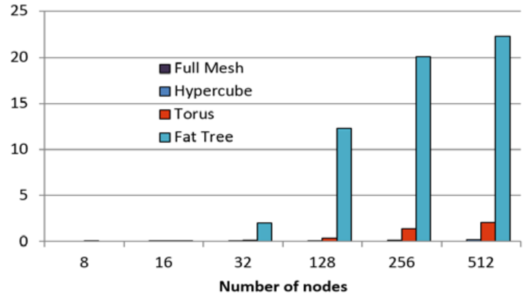


Fig. 7. Average jitter per flow

differences are relatively small between interconnections. By observing the evolution of this parameter, Figure 3(c), we note that the torus network can scale up and maintain a constant node degree. However, the fat tree is characterized by a higher increase rate in switch-to-switch port count when graphs scale.

The next section presents the results of simulating the expansion of the topologies with regard to key performance metrics: throughput, latency, lost rate and jitter.

III. Network Performance Simulations

The following simulations try to provide a comprehensive interpretation of the highly interconnectable networks assessed. In this scope, when setting up our simulation models, we are considering topology characteristics, communication pattern and the amount of data injected in the network as being the most relevant components for our scenarios.

In order to simulate the proposed scenarios, we used NS3, an open source discreet event network simulator capable of supporting network performance measurements related to throughput, delay, number of lost packets, jitter etc.

As in our previous evaluation [9], we have selected a shortest path based routing algorithm as opposed to a Spanning Tree Protocol implementation in order to evaluate the real potential of the topologies without blocking redundant ports for communication. The networks are subject to a uniformly distributed random traffic model that is widely accepted and is unbiased towards various topologies (e.g., fat tree behaves better under localized traffic patterns between neighbors in the same pod/cluster).

Besides the traffic pattern, the amount of traffic that will be injected in the interconnection network also plays an important role. Above a certain threshold the network begins to saturate, the ratio between throughput and injection rate starts to drop and the efficiency decreases. The saturation limit for hypercube and torus is 60% and 40%, respectively [15]. Therefore, we configure the input traffic level to be at 30% the maximum link capacity, below the saturation limit. Due to the random traffic pattern selection the results were averaged from a runtime of 30 seconds during which application based connection flows would be established between randomly paired hosts in the network.

As expected from the initial abstract metric analysis, we observe from Figure 4 that, even though the throughput per flow is declining when the networks scale from 8 nodes up to 512, the hypercube topology presents the most consistent behavior. A decrease of approximatively 5% in hypercube compared to 15% corresponding with the torus or a drastic 70% decrease in fat tree performance is measured.

Similar behavior occurs in the investigation of the delay (Figure 5) where even though the torus and hypercube demonstrate a similar delay, the 3-layer fat tree does not indicate well performance on scaling with an average of up to 40 ms per flow in 512 switched network. Same trend is observed when analyzing the rate of lost packets (Figure 6) and jitter (Figure 7) where hypercube and torus outperform the fat tree based network.

We have simulated and tested the performance of a full mesh topology (all switches have direct connections to every other switch in the network) under the same conditions as the other interconnects in order to perform a sanity check for the setup. First, this serves as a verification for all the configuration parameters employed, data rates, individual link delays, simulation time etc. Secondly, the test demonstrates the confidence in the results, which were not affected by hardware processing power when the topologies scaled from 8 to 512 nodes.

We can conclude from this section that, even though the cost indications and performance characteristics are similar for all studied topologies with nodes under a 64 count, clear superior operations are displayed by the hypercube followed by the torus.

IV. SDN For Data Center Interconects

Applying high radix topology designs into large data centers by employing current routing or switching technologies encounters numerous obstacles. A Layer 2 topology builds upon high density port switches that can process packets at line speed therefore, this implies less configuration and administration overhead. However, such a solution does not scale well due to the limitations of a flat topology and restrictions to a broadcast domain.

A routing based deployment can segment the broadcast domains, use existing performant routing protocols and present better scaling properties. Nevertheless, this comes at the expense of additional delays incurred by additional packet processing times in routers, which are not only slower but are also more expensive and require more complex administration [16]. In data centers, generally we see a compromise based on a combination of the two but also replacing network elements with expensive multilayer switches is an available option. However, SDN can truly get the best of both scenarios by giving routing capabilities to lower cost, white-box switches and automate administration operations with a topology manager module in the controller.

SDN adoption has raised many concerns about its impact on performance and scalability mainly due to the fundamental aspect of decoupling the control plane from the data plane. Ideas that a centralized controller is unlikely to scale as the network grows has led to some reluctance and certain expectations that some failure would occur when the number of incoming requests increases over supported limits. These assumptions can generally be sourced to the misconception that SDN implies the use of one physically centralized controller. Architectures involving a distributed control plane are, however, a valid way to construct a Software Defined Network and address the scalability issue, while also providing control plane resilience. Such solutions have already been demonstrated in projects like Onix and HyperFlow [17][18]. Yeganeh argues in [19], that there is no underlying bottleneck to SDN scalability, such a concern is not fundamentally unique to SDN. Even though a distributed SDN architecture would incur similar manageability problems as in a non-SDN approach, it would still be significantly easier to manage compared to having multiple heterogeneous switches running autonomous, vendor-specific applications.

V. Performance Evaluation of SDN DC Topologies

With the scope of evaluating and comparing SDN versus STP implementations on Hypercube, Torus and Jellyfish topologies, we used Mininet, a network emulation software that uses process-based virtualization to run multiple virtual OpenFlow switches and hosts on a same physical machine.

The SDN network controller of choice was Floodlight. The integrated topology and forwarding module enable the efficient use of the high number of redundant paths for networks like torus and hypercube and calculates connections based on shortest path between each pair of source and destination. However, link contention still occurs, influencing performance as seen in the results.

We have employed Iperf, a linux networking tool in order to measure network performance characteristics. We have focused on the same performance metrics identified as the most common attributes for network characterization in academic research, similar to our previous simulation scenarios: throughput, packet delay, jitter and loss rate. For the same reason as stated in Section 2, random data traffic patterns were configured, as well as a shortest path routing algorithm in the Floodlight SDN controller. To benchmark the fidelity of our test setup, a full mesh interconnection was assessed and subjected to the same tests.

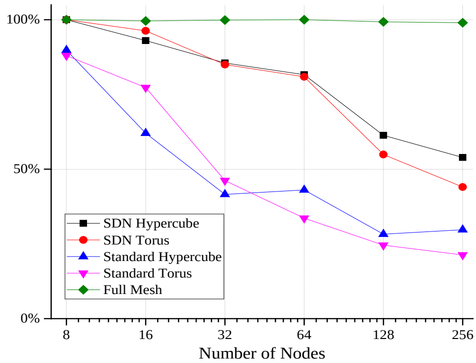


Fig. 8. Average throughput per node

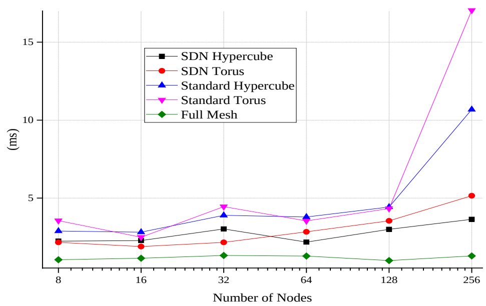


Fig. 9. Average latency per flow

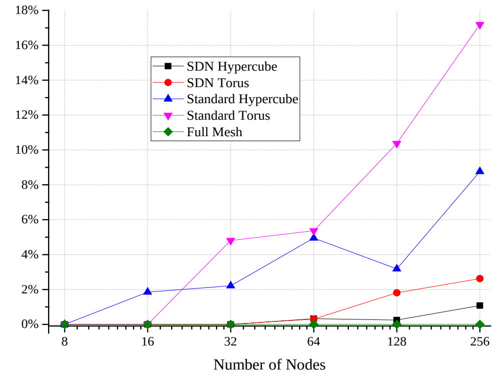


Fig. 10. Average loss rate per flow

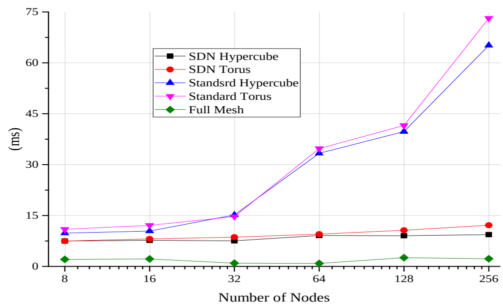


Fig. 11. Average jitter per flow

Normal traffic operation is presumed, as well as a constant injection rate (simple TCP transport) per application running on each server. Since network contention grows, the average throughput for each connection decrease when the networks expand in size including more switches and hosts.

In Figure 8, we observe the measured throughput is higher by roughly 45% for the SDN test cases with 256 nodes torus and hypercube. A 13ms decrease of packet delay, for SDN torus of the same size is noticeable in Figure 11. Loss rate is also reduced considerably for both topologies with at least 7 percentage points, as shown in Figure 9. Connection stability is improved with SDN technology by a reduction of jitter with 50ms, measurement that applies to networks of 256 switches; see Figure 10.

We plot in Figure 12 the comparison of the average throughput normalized by the theoretical link bandwidth capacity in the Jellyfish scenario. A two-fold increase in performance is observed for the SDN setup with 120 switches, with the difference slowly decreasing as the network scaled. The average packet delay, as seen in Figure 13, was considerably lower in the SDN scenario, with a small dependence on the number of switches employed and presenting less than 1/6th of the delay measured with STP for networks with more than 150 switches. Network jitter and loss rate were also more favorable in the SDN scenario: in Figure 14, we observed a reduction in jitter varying from 7% to 33%, and Figure 15 shows that the packet loss between the two systems remain within a 2% difference range independently of the number of switches.

VI. CONCLUSION AND FUTURE WORK

The first part of our paper focused on the rich and diversified design space of Data Center topologies and the differences between them. We can infer from our simulation results and the mathematical evaluation that, even though the cost indications and performance characteristics are similar for all studied topologies with under 64 nodes, clear superior operations are displayed by the hypercube followed by the torus on the downside of wiring complexity and scalability cost.

Our SDN versus STP emulation results demonstrate that the SDN deployments based on the studied topologies torus, hypercube and jellyfish, considerably outperform the STP implementation in throughput, latency, jitter and loss rate. No larger networks were tested due to the limitations of our emulation environment indicated by the degradation of the full mesh performance. Such results are representative for small to medium sized Data Centers however, much larger scales can be achieved with a distributed control plane solution [18][19], whereas the network setups discussed could represent individual clusters among many.

In addition to our previously presented work [11][12], concerning the performance gains achieved with SDN in

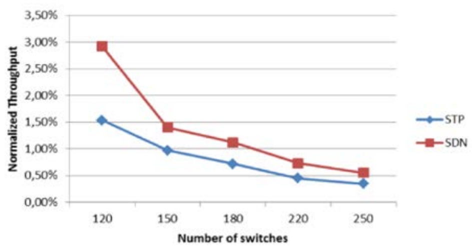


Fig. 12. Average throughput per node (Jellyfish)

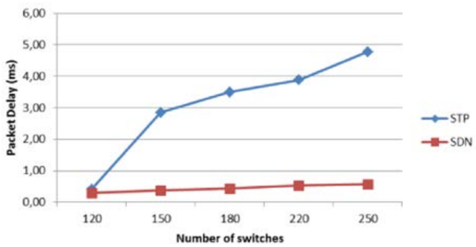


Fig. 13. Average delay per flow (Jellyfish)

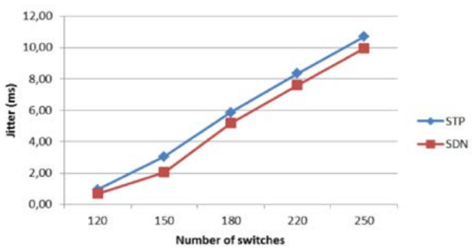


Fig. 14. Average jitter per flow (Jellyfish)

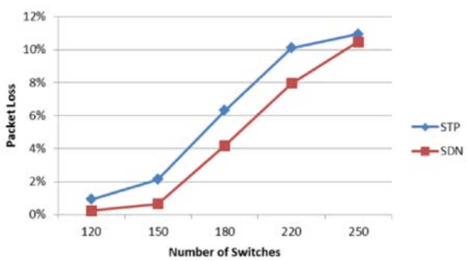


Fig. 15. Average loss rate per flow (Jellyfish)

highly interconnected network topologies, these results strengthen the arguments referring to how SDN can be used to leverage the multiple redundant paths in a network. Furthermore, we believe that a combination of SDN with MultiPath Transmission Control Protocol (MPTCP) can lead to an even more efficient network infrastructure utilization.

ACKNOWLEDGEMENTS

This work has been funded by the EU FP7 Marie Curie project ABACUS and by CNPq, *Conselho Nacional de Desenvolvimento Científico e Tecnológico* – Brazil.

REFERENCES

[1] Cisco Global Cloud Index: Forecast and Methodology, 2015–2020, White Paper, 2016.

[2] M. Al-Fares, et al., “A Scalable, Commodity Data Center Network Architecture,” in Proceedings of the ACM SIGCOMM conference on Data communication (ACM), pp. 63-74, New York, 2008.

[3] G. L. Vassoler, et al. “Twin Datacenter Interconnection Topology, " in IEEE Micro, Vol.34, Issue 5, pp.8-17, 2014.

[4] H. Wu, et al. "MDCube: a high performance network structure for modular data center interconnection, " In Proc. of ACM CoNEXT, 2009.

[5] A. Greenberg et al. ,“VL2: A scalable and flexible data center network, ” in Proc. of the ACM SIGCOMM conference on Datacommunication (ACM), pp.51-62, New York, 2009.

[6] J. Chen, P. Gillard, C. Li, “Performance Evaluation of Three Network-on-Chip Architectures (Invited),” in First IEEE International Conference on Communications in China: Communications QoS and Reliability (CQR), pp. 91-96, 2012.

[7] J. K. Dennis Abts, "High Performance Datacenter Networks", by Morgan & Claypool, 2011.

[8] A. Singla et al., "Jellyfish: networking data centers randomly" in Proc. Of the 9th USENIX conference on Networked Systems Design and Implementation, pp. 17-17, California, 2012.

[9] B. Andrus, J. J. V. Olmos and I. T. Monroy, ”Performance Evaluation of Two Highly Interconnected Data Center Networks”, in Proc. International Conference on Transparent Optical Networks, Budapest, 2015.

[10] Network Simulator 3, found at: https://www.nsnam.org/.

[11] B. Andrus, J. J. V. Olmos, V. Mehmeri and I. T. Monroy „SDN data center performance evaluation of torus and hypercube interconnecting schemes, “ in Advances in Wireless and Optical Communications (RTUWO), 2015.

[12] V. Mehmeri, J. J. V. Olmos, and I. T. Monroy “Capacity Extension of Software Defined Data Center Networks With Jellyfish Topology, ” in Asia Communications and Photonics Conference (ACP), 2015.

[13] A. Singla et al.„ "Jellyfish: networking data centers randomly" in Proceedings Of the 9th USENIX conference on Networked Systems Design and Implementation, California, pp. 17-17, 2012.

[14] W. J. Dally, “Performance Analysis Of K-Ary N-Cube Interconnection Networks,” in IEEE Transactions On Computers, vol. 39, no. 6, pp. 775-785, 1990.

[15] A. S. Muhammed Mudawwar, "The k-ary n-cube Network and its Dual: a Comparative Study," in IASTED International Conference on Parallel and Distributed Computing and Systems, 2001.

[16] K. Jayaswal, “Administering Data Centers”, Wiley Publishing, 2005.

[17] T. Koponen et al., “Onix: A Distributed Control Platform for Large-scale Production Networks”, in Proc. of the 9th USENIX conference on Operating systems design and implementation, California, Article No. 1-6, 2010.

[18] A. Tootoonchian and Y. Ganjali, “HyperFlow: a Distributed Control Plane for OpenFlow”, in Proc. of the Internet Network Management Conference on Research on Enterprise Networking, California, pp.3-3, 2010.

[19] S.H Yeganeh, et al., "On scalability of software-defined networking", in IEEE Communications Magazine (Institute of Electrical and Electronics Engineers, New York), Vol. 51, Issue 2, 2013.

Paper G: Evaluation and Experimental Demonstration of SDN-Enabled Flexi-grid Optical Domain Controller based on NETCONF/YANG

B. Andrus, A. Autenrieth, T. Szyrkowiec, J.J. V. Olmos, and I. T. Monroy "Evaluation and Experimental Demonstration of SDN-Enabled Flexi-grid Optical Domain Controller based on NETCONF/YANG," Accepted for publication *in Network Operations and Management Symposium (NOMS)*, Taipei, 2018.

Evaluation and Experimental Demonstration of SDN-Enabled Flexi-grid Optical Domain Controller based on NETCONF/YANG

B. Andrus(1,2), A. Autenrieth(1), T. Szyrkowiec(1), J.J. Vegas Olmos (2,3), I. Tafur Monroy(2,4)

(1)ADVA Optical Networking, Martinsried, Germany (bandrus@advaoptical.com)
(2)Technical University of Denmark, Department of Photonics Engineering, Lyngby, Denmark
(3)Mellanox Technologies, Roskilde, Denmark
(4)Technical University of Eindhoven, Department of Electrical Engineering, Netherlands

***Abstract* – Flexible spectrum assignment in Elastic Optical Networks (EON) has emerged as a potential solution for allowing dynamic and elastic management of available bandwidth resources. In this paper, we demonstrate and evaluate our developed flexi-grid optical domain controller based on NETCONF/YANG. Our proposed modular architecture, based on Finite State Machines (FSMs), allows the flexibility to deploy the controller either in a centralized or in a distributed state for on the fly encrypted device management connections. A testbed composed of two physical Sliceable Bandwidth Variable Transponders (SBVTs) and an emulated flexi-grid optical network was used for our software evaluation. Controller startup and synchronization time, as well as media channel setup time are evaluated to compare the two deployment options and assess network scaling effects. Results demonstrate that our software is scalable by maintaining a relatively constant startup time on the networks tested (i.e., 1 to 64 nodes) in both deployment options. Software scalability is also supported by the media channel setup time, which presents a modest log scale growth when increasing the number of nodes from one to 64.**

***Index Terms* – SDN; flexi-grid; Elastic Optical Networks; WDM; NETCONF; YANG; encrypted communication.**

I. INTRODUCTION

Optical Wavelength-Division Multiplexing (WDM) has provided increased transport capacity using parallel wavelength channels. However, overall optical spectrum utilization is inefficient due to its fixed channel spacing grid and standard rate transmission. Introducing flexibility into the spectrum assignment can alleviate these limitations by providing variable fine-grained spectrum slots (e.g., 12.5 GHz vs. 50 GHz or 100 GHz) [1].

As service demands continue to grow, wide area networks, including transport interconnects, need to adapt to offer their resources on demand [2]. However, optical networks are evolving in heterogeneous multi-domain, multi-vendor and multi-technology deployments, which poses new interconnection problems [3],[4].

Provisioning paths across multi-domain optical networks has already been defined as a challenge [4], which is amplified by the deployment of heterogeneous devices. Conventional Network Management Systems are unsuited to handle multivendor network devices due to their proprietary and closed interfaces. On the other hand, SDN architectural principles offer a variety of possibilities when looking to plan, control, and manage multi-vendor network resources sometimes at the cost of scalability[5],[6] .

NETCONF protocol [7] is an SDN enabler that offers customizable control and management capabilities. NETCONF, together with the YANG [8] (data modeling language), have been regarded with increasing interest by network operators because of the possibility to standardize common models for configuration and management data in a vendor-neutral way.

Previous research [9] has shown that multi-domain optical controllers using OpenFlow (OF) agent deployments based on NOX[10] (developed in C++) can significantly improve path setup times compared to conventional GMPLS-based control plane utilized in optical networks. However, decreasing the path

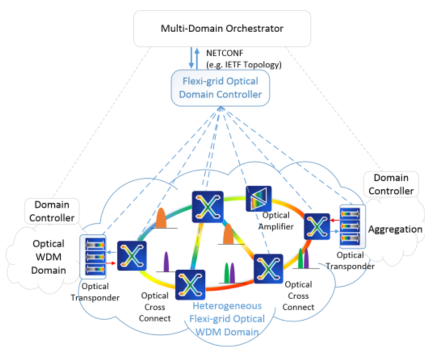


Fig. 1. Flexi-grid optical domain controller with a NETCONF northbound interface based on IETF network topology model.

978-1-5386-3416-5/18/$31.00 ©2018 IEEE

setup time even more can be accomplished and additional scalability effects on the controller performance can be evaluated further.

With this work, we aim to demonstrate a centralized, scalable and versatile control plane solution for heterogeneous flexi-grid optical networks. By introducing a NETCONF programmatic interface based on IETF optical flexi-grid YANG schemas [11], we plan to provide a tool for end-to-end path provisioning for flexi-grid networks, vital for an optical multi-domain orchestration use case, as seen in Fig.1.

This report is structured as follows: Section II offers an introduction to flexi-grid optical networks while Section III presents data modeling considerations for optical transport networks. Section IV describes our software implementation and Section V details the demonstration setups. In Section VI, we report on the functionality and evaluation of our developed controller. The last chapter summarizes the conclusion and lessons learned from our work.

II. Flexible WDM Grid Optical Networks

Conventional fiber spectrum is normally segmented using wavelength division multiplexing into separate channels with 50 or 100 GHz spacing. The wavelengths transmitted over each channel are generated by transponders with fixed line rates (e.g., 10, 40, 100 Gb/s). To support higher bit rates over the same channel, especially beyond 400 Gb/s, a finer granularity for grid assignment is defined by flexi-grid specifications [1] allowing concatenation of adjacent channels to form super channels (i.e., several optical carriers combined to create a channel of desired capacity).

Flexi-grid refers to the renewed set of nominal central frequencies with narrower channel spacing and a more efficient optical spectrum management. In the context of flexi-grid, a frequency slot is a logical abstraction that represents a frequency range defined by a nominal central frequency ($f = 193.1\ THz + n \times 6.25\ GHz$) and its slot width ($SW = m \times 12.5\ GHz$). In the view of Optical Transport Networks (OTN) [12], the optical signal is guided to its destination transponder using media channels. A media channel

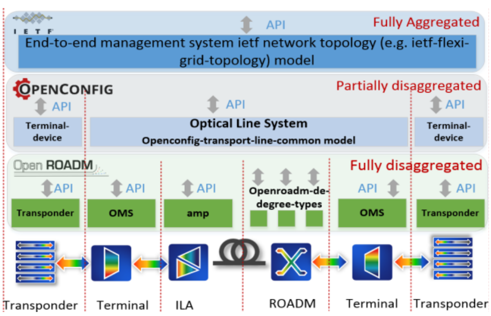


Fig. 2. Comparison of aggregated and disaggregated network models. (ILA - Amplifier; OMS – Optical Multiplex System)

defines the path through the network between end transponders (i.e., intermediary nodes and links) as well as the frequency slots it occupies on each individual link.

Wavelength Switched Optical Networks (WSONs) are WDM networks that include switching elements. In WSONs, WDM channels are utilized to interconnect IP capable devices through multi-degree Reconfigurable optical add-drop multiplexers (ROADM) to overcome hop-by-hop optoelectronic processing. Switching on this level is done based on frequency slots.

III. Information Models For Optical Transport Networks

YANG data models provide semantic-rich descriptions of the data structure representations of network resources and services. The YANG abstraction layers can create an imposed agreement not only the between SDN control system and the network elements but also between the system and the northbound application depending on the network architecture (e.g., aggregate or disaggregate).

On the one hand, an aggregated model, where the entire network acts as a single managed system, is optimized for network operation and can yield higher performance using vendor specific interfaces. End-to-end service orchestration is simplified reducing overall management processes. However, such an approach can affect innovation due to differences in the life cycle of various network elements and also because of strong component interdependency.

On the other hand, in a disaggregated [13] network model each component is viewed and designed as an individual network element. Advantages that arise from such an approach focus mainly on cost efficiency and innovation optimization. However, operating a disaggregated system not only complicates the central control architecture but also increases the overall synchronization challenge. As a result, end-to-end service orchestration reaches new complexity levels.

The lack of standard interfaces for network automation has lead various work groups and standards organizations to work together to define common configuration data models able to support multivendor network management. Recent initiatives such as OpenConfig, OpenRoadm, ONF or IETF CCAMP provide steps in the right direction with respect to optical interoperability and unified management.

OpenROADM [14] focuses on opening up traditionally proprietary ROADMs for SDN development. The solution adopted by OpenROADM in the first release (v.1) follows a fully distributed approach in which every component inside the device is modeled individually, in detail, from tunable transponders to ROADM components (e.g., pluggable optics, degrees, ODU/OTN/multiplex interfaces etc.) as seen in Fig. 2. End-to-end service paths are also supported in the form of Remote Procedure Calls (RPC).

OpenConfig [15] follows a similar goal of compiling a consistent set of vendor-neutral data models. Their approach differs however, by adopting a partially distributed network model as seen in Fig. 2.

Due to rapid rate of innovation in coherent DWDM transponders, a solution to decouple the transponder from the rest of the optical line system allows the network operator to take advantage of the best transponder at any given time.

First, elements comprising an optical transport line system (e.g., Amplifiers, ROADMs) are grouped and defined in a yang module (transport-line-common.yang). The Optical Line System (OLS) module defines the ROADMs as configurable switching elements with input and output ports.

Secondly, OpenConfig describes an optical terminal device for DWDM transport that offers support for optical channel configuration such as wavelength/frequency, power, Bit Error Rate (BER) and operational mode. However, in the current released version, the models lack support for flexible channel spacing assignment and an extension in this direction would not be straight forward since the current optical channel definition allows only one frequency assignment.

Other works, like [16], focus on creating new open models that focus on specific parts of an Elastic Optical Network. In [16] and [17], the authors offer details on the development of a comprehensive control system for flexible transponders including parameters like adaptable modulation, rate, FEC, slice-ability and monitoring of pre-FEC BER which are used to automatically adapt the FEC and the baud rate to maintain a robust communication transmission with the help of a simple two-state FSM.

The IETF CCAMP working group has released a data model that represents, retrieves and manipulates elements from a flexi-grid capable optical network as a whole [11]. IETF focuses on an aggregated approach providing an overview of the optical network topology combined with the underlying physical layer.

On the one hand, an optical Traffic Engineering Database (TED) defines the main elements. Flexi-grid nodes (abstract ROADM model with input/output ports and internal port connectivity matrix), transponders (tunnel termination points with variable FEC and modulation) and links (connections with source and destination nodes and ports with additional flexi-grid attributes like base frequency, maximum nominal central frequency, allowed frequency granularity and slot width granularity).

On the other hand, the model also describes media channel structures containing the intermediate nodes and links referenced in the optical TED as well as the frequency slot.

Since this presents a more abstract view of the network topology, any other physical layer parameters regarding optical channel interfaces for DWDM applications are tackled and defined in a different project [18].

IV. Software Implementation

Due to its fundamental scope of supporting flexi-grid enabled optical networks and its expandable TED-oriented information structure, we have chosen to implement our controller software based on the IETF flexi-grid model[11]. Fig. 3 displays the architecture of our developed software. To achieve a flexible and modular design we have adopted Akka [19] a java toolkit. Its actor/module-based model facilitates the development of parallel and non-blocking software applications. In our work, we have assigned each actor (fundamental unit of computation) very specific roles with the scope of fulfilling the functionality requirements of each module (i.e., Network Driver, Device Driver etc.).

Communication between modules is asynchronous and occurs through an exchange of self-defined immutable messages (e.g., NetConfMsg, DevConfMsg etc.), which contain information like type of message, network node ID, configuration parameters and corresponding values etc.

Two deployment scenarios are envisioned taking advantage of the modular nature of our controller. On the one hand, a centralized version is foreseen, in which all modules run locally, presumably in a central office, having proprietary connections to each network device through (i.e., custom API), possibly unencrypted. As a brownfield solution, we target a distributed setup, in which Device Driver modules are spawned in the same location as the physical devices, (i.e., Point of Presence – PoP – hosting the network ROADMs). The major advantage from this approach would be the SSL encryption provided by the Akka underlying remote message exchange that establishes a secure management connection between the central office and the remote PoPs.

To cover all possible device and network configuration scenarios and to avoid application deadlocks, we have developed the modules with an

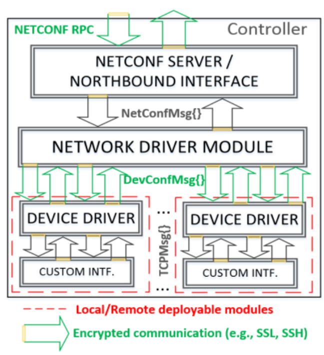


Fig. 3. Modular architecture of flexi-grid network controller. Device Drivers can be deployed remotely, providing on the fly encryption for Device Configuration messages (DevConfMsg)

architecture based on Finite State Machines (FSMs). Simplified representations of the NetDriver and DevDriver FSMs are described in Fig. 4(b) and 4(c). The FSM models allow us to define the asynchronous communication between the modules.

A. NETCONF Server

As a northbound API, which allows the software to provide a unified view of the flexi-grid network topology to a multi-domain orchestrator, we have built a NETCONF server interface based on netconf4J java libraries [20].

A NETCONF processing factory not only categorizes incoming messages based on XML tags (e.g., hello, get-config, edit-config etc.) but also identifies the configuration elements (e.g., network, nodes, media-channel etc.). Deserialization is accomplished by a set of JavaBeans generated from the original IETF YANG models [11], which performs a validation of the XML data against the YANG schemas. The JavaBeans are generated using an in-house developed plugin for the Pyang tool [21]. Deserialization errors return a NETCONF <rpc-error> with an "Invalid Data" tag. In the current implementation, all other configuration errors are treated equally and tagged as "Configuration Error".

B. Network Driver Module

The Network Driver module, detailed in Fig. 4(b), oversees parsing configuration and operational information for all the devices managed by the application. According to the architecture, during initialization, the Network Driver reads the topology information (node-IDs, number of ports and optical links) from a JSON configuration file and creates individual Device Driver for each network device.

Next, the Network Driver updates the internal network state by reading the supported configuration from all the connected devices. After synchronization, the module transitions to an Idle State awaiting NETCONF messages. From this state it can either go to a Node Configuration State, which is specific for TED defined parameters (e.g., modulation, FEC, internal connections etc.) or in a Media Channel Configuration State for optical channel defined parameters (e.g., effective frequency slot, source/destination nodes, transponders etc.).

The Flow Diagram in Fig. 4(a) details the message exchange that occurs between the different application modules during a startup stage and a generic NETCONF edit-config operation (i.e., node or media channel config).

When entering the Media Channel (MC) Configuration State, the received MC configuration request message is split into individual device configuration messages and sent to each individual Device Driver. When all the replies have been received, the MC is stored locally and a MC creation acknowledgement is sent back to the NETCONF interface before transitioning back to the Idle State.

For any edit-config operations received, related to network nodes, the Network Driver enters a Node Configuration State. When entering this state, the edit-config message is parsed and sent to the appropriate Device Driver. Upon receiving the device reply message, the internal network configuration is updated and a write acknowledgement is sent back to the NETCONF interface.

When receiving NETCONF get-config operations, the Network Driver queries the internal state of the network and sends a copy of the stored information without having to transition from the Idle State or read the device configuration again.

C. Device Driver Module

The role of the Device Driver, as seen in Fig. 4(c), is to establish a connection to the physical device using its specific Interface Module (e.g., TCP Socket). In addition, it is also responsible for the translation between the device specific configuration API and the flexi-grid based data structures.

At initialization, the Device Driver establishes a

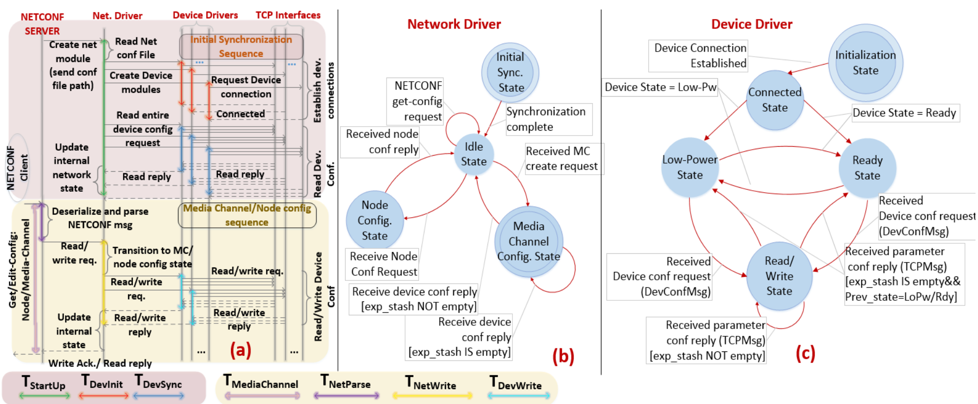


Fig. 4. (**a**) Flexi-grid domain controller flow diagram (e.g., synchronization and MC/Node configuration sequence); FSM-based Network Driver (**b**) and Device Driver (**c**); (MC-Media Channel, FSM – Finite State Machine).

connection to the physical device. After this, the module changes state from Idle to Connected and proceeds to query the device for the running state (e.g., Low Power or Ready). The transition to the next FSM state is determined by the reply received. If the device is ready to transmit data then the module shifts into a Ready state; if the device is in a powered down administrative state where fundamental parameters like modulation format, FEC etc. can be reconfigured, then the Driver assumes the Low Power state.

When a read/write request is received from the Network module, the Device Driver transitions into a Read/Write state. The individual configuration parameters are extracted and read/written from/to the device sequentially with the help of the Custom API module. When the final parameter has been configured, the module transitions back to the previous state: Low Power or Ready.

V. Demonstration Setup

We have accomplished the validation and evaluation of the two deployment variants of our controller application by using a testbed comprised of a pair of physical SBVTs, a flexi-grid network emulator and a test NETCONF client script. In the centralized demonstration setup, the controller runs on a local PC and the emulator run on separate remote PC (both configured with 2 core and 4 thread CPUs @2.1GHz, 8 GB RAM and 1Gb/s network connections to the setup). In the distributed option, displayed in Fig. 5(a), the Device Driver agents are initialized in independent JVMs hosted on the remote JVM machine 'closer' to the emulated nodes (ROADMs).

We have designed a basic flexi-grid virtual network with the scope of emulating flexi-grid ROADMs and links in separate structures. The emulator consists of basic modules representing flexi-grid ROADMs, which parse the controller device configuration messages and read/write to individual configuration files.

VI. Experimental Demonstration

First, we evaluate the initial network synchronization sequence – $T_{StartUp}$. Due to Akka's migration capability, which can re-spawn an entire application on a remote machine, we consider the startup time to be a relevant metric for various backup and failure recovery scenarios. Second, we asses a media channel configuration (path setup time), - $T_{MediaChannel}$.

As opposed to [9], where a NOX based OF optical controller achieves path setup time in over 10 s across a fixed network of 3 physical ROADMs, we consider our metrics relevant to the performance characterization of our software under network scaling conditions (i.e., up to 64 emulated nodes) in both a centralized and distributed deployment. A representation of $T_{StartUp}$ and $T_{MediaChannel}$ and their components was shown in Fig. 4(a).

A. Startup time - $T_{StartUp}$

In Fig. 5(b) we see highlighted a serialized encoding of a DevConfMsg containing parameters and their corresponding options (e.g., Modulation, FEC, Channel number etc.). The initial deployment of a Device Driver and its corresponding Custom API module *($T_{DevInit}$)*, which includes the process of establishing the connection with the physical device, is a major step in the startup procedure. The next step is retrieving the running configuration and establishing a Device Driver level synchronization *($T_{DevSync}$)*. After receiving, parsing and storing the configuration data from all Device Drivers, the Network Driver synchronization is complete *($T_{StartUp}$)* and the application is ready to receive NETCONF configuration requests.

The main delay component $T_{DevInit}$, takes longer in the distributed configuration due to remote deployment and secure SSL connection establishment

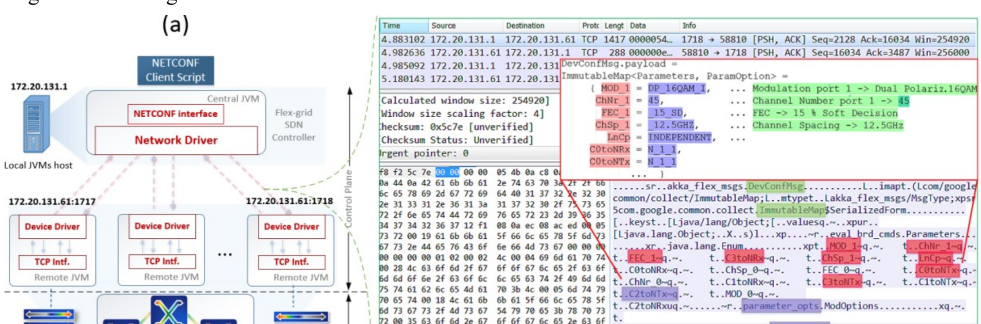


Fig. 5. (a) Demonstration setup (distributed deployment), (b) Device Configuration Message capture

with the Device agents, as seen in Fig. 6(a). However, due to the multithreading architecture of our software this does not have a cumulative effect on network scaling and the final synchronization time retains a relatively constant value (i.e., 920 ms for the centralized and 1150 ms for the distributed) with slight variations likely due to garbage collection cycles in the JVM. We point out that even though the device synchronization is faster in the distributed version - Fig. 6(b), the component ratio of $T_{DevSync}$ from the total $T_{StartUp}$ is much smaller than $T_{DevInit}$ and therefore the final $T_{StartUp}$ is lower when the software is centralized supported by Fig. 6(c).

B. Media channel configuration time – $T_{MediaChannel}$

We select for testing a media channel that connects the two end transponders and occupies the frequency slot <*n=7, m=3*> (f=193.14375 THz, SW = 37.5GHz). The number of links in the media channel for each network size test is equal to the number of nodes – 1.

During the media channel configuration, we determine a few key processes. When receiving the initial NETCONF message the data is deserialized and split for individual Device Driver configuration ($T_{NetParse}$). A per Device Driver configuration is then assessed ($T_{DevWrite}$) where the driver writes out the new configuration to the physical device. Additionally, writing and then storing the configuration data from all Device Drivers at a network level ($T_{NetWrite}$) is evaluated as well. The confirmation received back at the test application marks the end of the media channel configuration time ($T_{MediaChannel}$).

Our results reveal that the centralized version has

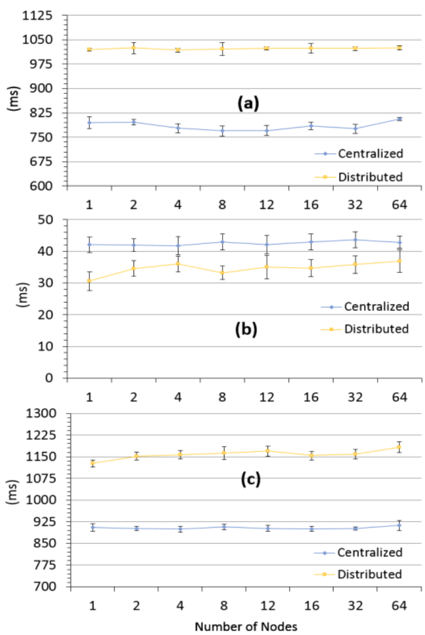


Fig. 6. (a) Device Driver Initialization - $T_{DevInit}$, (b) Device Driver Synchronization - $T_{DevSync}$, (c) Total Startup Time - $T_{StartUp}$

a faster configuration time than the distributed in all tested network sizes - Fig. 7. The main reason is the additional software encryption and decryption of DevConMsgs between the remote modules.

By scaling the network, we see a boost in the configuration delay (i.e., a logarithmic scale growth) due to an increase in the number of nodes supported and in the number of message exchanges at the Network Driver level. However, the effect of $T_{DevWrite}$ is still not cumulative on the $T_{NetWrite}$. As expected, the deployment choice does not have an influence on $T_{NetParse}$ which remains similar in both conditions and is only a small component in $T_{MediaChannel}$. Therefore, on a network with 64 nodes the centralized controller has a $T_{MediaChannel}$ of 101 ms and the distributed 182 ms.

VII. CONCLUSION

We conclude that the IETF YANG schema is mostly tailored towards network configuration however not towards monitoring as pre-FEC BER is missing. Additional traffic engineering configuration and monitoring capabilities (e.g., delay, bandwidth metrics, link statistics etc.) are provided in the base models imported and augmented by the flexi-grid YANG. However, these dependencies between hierarchical models translate into additional mandatory processing overhead regardless if they are needed or not for a specific use case.

Based on our test results conducted in a lab environment, with minimum latency on the control plane path between network nodes and the controller, we argue that both centralized and distributed approaches demonstrate relatively small influences on performance when startup and synchronization occurs. However, a modest logarithmic-like scale increase in delay for media channel setup is observed on network scaling. The modular, FSM-based multi-thread implementation presents good scaling capabilities. In both tests, the centralized version presents lower startup and configuration times due to overhead in the distributed deployment caused by additional remote agent initialization procedures but mostly by encryption.

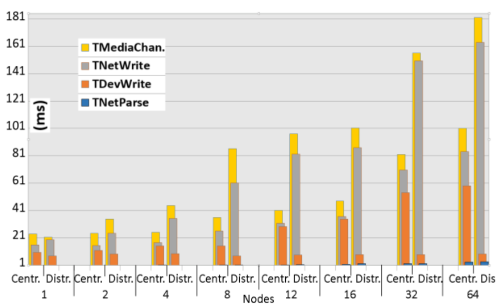


Fig. 7. End-to-End media channel configuration steps ($T_{MediaChannel}$): NETCONF message parse time $T_{NetParse}$, Device Driver write time $T_{DevWrite}$, Network Driver Write Time ($T_{NetWrite}$)

ACKNOWLEDGEMENT

This work was funded by EU Projects ABACUS and Metro-Haul.

REFERENCES

[1] ITU-T, "Spectral grids for WDM applications: DWDM frequency grid ", Recommendation G.6941, February 2012.
[2] V. S. Shukla, “SDN Transport Architecture and Challenges,” OFC2015, W4J.1, 2015.
[3] M. Chamania and A. Jukan, “A survey of inter-domain peering and provisioning solutions for the next generation optical networks,” IEEE Commun. Surv. Tutorials, vol. 11, no. 1, pp. 33–51, 2009.
[4] . K. Liang, M. Rahnamay-Naeini, H. M. K. Alazemi, N. MinAllah, M. Peng, and N. Ghani, “Optimization approach for multi-domain optical network provisioning,” J. Opt. Commun. Netw., vol. 5, no. 5, pp. 1413–1424, 2013.
[5] Z. Qingyun, C. Ming, D. Ke, X. Bo "On Generality of the Data Plane and Scalability of the Control Plane in Software-Defined Networking," in China Communications, 2014.
[6] M. Karakus, A. Durresi " A Scalability Metric for Control Planes in Software Defined Networks (SDNs)," in Advanced Information Networking and Applications (AINA), 2016
[7] M. Bjorklund, et al., "Network Configuration Protocol (NETCONF)," RFC 6241, June 2011, found at: < https://tools.ietf.org/html/rfc6241>.
[8] M. Bjorklund, et al, "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, < http://www.rfc-editor.org/info/rfc6020>.
[9] M. Channegowda, R. Nejabati, et al. "Experimental demonstration of an OpenFlow based software-defined optical network employing packet, fixed and flexible DWDM grid technologies on an international multidomain testbed," in Optics Express, Vol. 21, Issue 5, pp. 5487-5498, (2013).
[10] NOX, SDN controller found at: < https://github.com/noxrepo/nox>.
[11] J. Vergara, et al., "YANG data model for Flexi-Grid Optical Networks", IETF draft-vergara-flexigrid-yang-02, Internet Draft, March. 2017.
[12] ITU-T Recommendation G.872: “Architecture of optical transport networks”, January 2017.
[13] P. Sterling, "Bringing Disaggregation to Transport Networks", Fujitsu, White Paper, October 2015.
[14] OpenROADM, "Metro Open ROADM Network Model", White Paper, Version 1.1., 21st July 2016.
[15] R. Shakir, et al., “Consistent Modeling of Operational State Data in YANG”, IETF draft-openconfig-netmod-opstate-01, 2015.
[16] M. Dallaglio, N. Sambo, F. Cugini and P. Castoldi "YANG Models for Vendor-Neutral Optical Networks, Reconfigurable through State Machine," in IEEE Communications Magazine, Volume: 55, Issue: 8, pp. 170 - 178, 2017.
[17] M. Dallaglio, N. Sambo, J. Akhtar, F. Cugini and P. Castoldi, "YANG model and NETCONF protocol for control and management of elastic optical networks", Optical Fiber Communications Conference and Exhibition (OFC), 2016.
[18] G. Galimberti, et. al. "A YANG model to manage optical interface parameters of DWDM", IETF draft-dharini-netmod-g-698-2-yang-00, July 2014.
[19] Akka toolkit, found at: < http://akka.io/>
[20] NETCONF java libraries, netconf4j found at: < https://github.com/dana-i2cat/netconf4j>
[21] Pyang, An extensible YANG validator and converter in python, found at: < https://github.com/mbj4668/pyang>

Bibliography

[1] Cisco, "The Zettabyte Era: Trends and Analysis June 2017", White Paper, June 2017.

[2] S. Pachnicke, et al., "The Frontiers of Optical Access Networks", International Conference on Optical Network Design and Modeling (ONDM), Pisa, Italy, May 2015

[3] D. Breuer et al., "Fixed-mobile convergence Introduction: Historical overview", (Tutorial) International Conference on High Performance Switching and Routing (HPSR) , found at: `http://www.ict-combo.eu/index.php?id=fmc-tutorial-2015`, 2015.

[4] ITU "Fixed-Mobile Convergence," Discussion Paper, Global Symposium for Regulations, Dubai 2007.

[5] COMBO (COnvergence of fixed and Mobile BrOadband access/aggregation networks) project, European Community's Seventh Framework Program FP7-2007-2013 ICT Integrated Project, under grant agreement 317762, `http://www.ict-combo.eu/`.

[6] China Mobile Research Institute, "C-RAN: The Road Towards Green RAN", White Paper, 2011.

[7] D. Chitimalla, K. Kondepu, L. Valcarenghi, M. Tornatore, B. Mukherjee "5G Fronthaul-Latency and Jitter Studies of CPRI Over Ethernet," in Journal of Optical Communications and Networking (JOCN), Volume: 9, Issue: 2, Feb. 2017.

[8] NGMN Alliance, "Future Study on Critical C-RAN Technologies," A Deliverable by the NGMN Alliance, White Paper, 31st of March 2015.

[9] J. S. Wey, et al, "Physical Layer Aspects of NG-PON2 Standards - Part1: Optical Link Design [Invited]", OSA J. Opt. Comm. Netw., vol. 8, no. 1, pp. 33-42, January 2016.

[10] Y. Luo, et al, "Physical Layer Aspects of NG-PON2 Standards - Part2: System Design and Technology Feasibility [Invited]", OSA J. Opt. Comm. Netw., vol. 8, no. 1, pp. 43-52, January 2016.

[11] Z. Tayq, et al, "Fronthaul Performance Demonstration in a WDM-PON-Based Convergent Network", European Conference on Networks and Communications (EuCNC), Athens, Greece, June 2016.

[12] A. Zakrzewska, S. Ruepp, M. S. Berger "Towards converged 5G mobile networks-challenges and current trends," in ITU Kaleidoscope Academic Conference: Living in a converged world - Impossible without standards?, 2014

[13] T. Mamouni, J. Torrijos Gijón, P. Olaszi, X. Lagrange "Universal AAA for hybrid accesses," in EuCNC 2015

[14] Younes Khadraoui; Xavier Lagrange; Stefan Höst; Thomas Monath, "On connection control and traffic optimisation in FMC networks",

[15] COMBO Alliance, "Analysis of horizontal targets for functional convergence", EU ICT-COMBO project deliverable 3.2, pp 45 - 69,

[16] Erik Nygren, et al., "The Akamai network: a platform for highperformance internet applications". SIGOPS Operating Systems Review Vol. 44, No. 3, pp. 2-19, August 2010 (DOI: 10.1145/1842733.1842736).

[17] Claudio Imbrenda, et al., "Analyzing cacheability in the access network with HACkSAw.", International Conference on Information-Centric Networking (ICN '14), Paris, France September 2014 (DOI: 10.1145/2660129.2660135).

[18] Salah-Eddine Elayoubi and James Roberts. "Performance and Cost Effectiveness of Caching in Mobile Access Networks", ICN '15, San Francisco, USA, September 2015 (DOI: 10.1145/2810156.2810168).

[19] Zhe Li, Jean-Charles Point, Selami Ciftci, Onur Eker, Giulia Mauri, Marco Savi, Giacomo Verticale, "ICN Based Shared Caching in Future Converged Fixed and Mobile Network", International Conference on High Performance Switching and

Routing (HPSR), 2015. http://www.ictcombo.eu/data/uploads/deliverables/combo_d3.2_v2.0_pu.pdf, 30 April 2015. International Conference on High Performance Switching and Routing (HPSR), 2015

[20] R. Kohavi, R. M. Henne, and D. Sommerfield. "Practical Guide to Controlled Experiments on the Web: Listen to Your Customers not to the HiPPO," KDD, 2007.

[21] Heavy Reading (on behalf of Juniper Networks and Affirmed Networks) "Designing 5G-Ready Mobile Core Networks" White Paper found at:https://www.affirmednetworks.com/wp-content/themes/renden/pdf/5G_Whitepaper_Heavy_Reading.pdf

[22] Huawei, "5G: A Technology Vision," White Paper, February 2014.

[23] K. Church, J. Hamilton, A. Greenberg, "On delivering embarrassingly distributed cloud services," in Hot Topics in Networks (HotNets) workshops, October 2008.

[24] Cisco "Cisco Global Cloud Index: Forecast and Methodology, 2015-2020," White Paper, 2016

[25] J. Kim, D. Abts, "High Performance Datacenter Network: Architectures, Algorithms and Opportunities", Morgan & Claypool, 2011.

[26] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel, "The cost of a cloud: Research problems in data center networks," SIGCOMM Computer Communication Review, vol. 39, no. 1, pp. 68-73, Jan. 2009.

[27] ITU Recommendation G694.1 "Spectral grids for WDM applications: DWDM frequency grid", 2014.

[28] Víctor López, Luis Velasco "Elastic Optical Networks: Architectures, Technologies, and Control", published by Springer, 2016

[29] R. Rumipamba-Zambrano, J. Perelló, A. Pagés, J. M. Gené, and S. Spadaro "Influence of the Spatial Super Channel Guard-Band Width on the Performance of Dynamic Flex-Grid/SDM Optical Core Networks ," in International Conference on Transparent Optical Networks (ICTON), 2016

[30] M. Bjorklund, et al, "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, http://www.rfc-editor.org/info/rfc6020.

[31] A. Clemm, J. Medved, R. Varga, N. Bahadur, H. Ananthakrishnan and X. Liu "A Data Model for Network Topologies, ", IETF draft-ietf-i2rs-yang-network-topo-16.txt, Internet Draft, September 2017.

[32] P. Sterling, "Bringing Disaggregation to Transport Networks", Fujitsu, White Paper, October 2015.

[33] OpenROADM, "Metro Open ROADM NETWORK MODEL", White Paper, Version 1.1., 21st July 2016.

[34] OpenROADM MSA Specifications found at: http://0201.nccdn.net/1_2/000/000/162/00c/20160510-Open-ROADM-MSA-specification-ver-1-03.xlsx.

[35] White Paper "Open ROADM overview", OpenROADM MSA, Version 1.0, October 2016.

[36] R. Shakir, "Consistent Modeling of Operational State Data in YANG", IETF draft-openconfig-netmod-opstate-01, 2015.

[37] J. Vergara, et al.,"YANG data model for Flexi-Grid Optical Networks", IETF draft-vergara-flexigrid-yang-02, Internet Draft, March. 2017.

[38] O. Gonzalez de Dios et. al., "Framework and Requirements for GMPLS-Based Control of Flexi-Grid Dense Wavelength Division Multiplexing (DWDM) Networks", RFC7698, ISSN 2070-1721, November 2015.

[39] G.Galimberti, et. al. "A YANG model to manage t optical interface parameters of DWDM", IETF draft-dharini-netmod-g-698-2-yang-00, July 2014.

[40] D. Lee, J. Bang, Y. Ran Cho, and H. Jung, "Integrated Management Architecture for FMC Based on IMS/NGN, ", in International Conference on Next Generation Networks, 2006

[41] Hemant Borale; Vaibhav Mishra; P. W. Wani, "Fixed and Mobile Convergence: A new approach, " in International Conference on Next Generation Networks, 2010.

[42] A. Yarali, A. Cherry, and B. Ahsant "FMC: The Convergence of Networks, Devices and Services, " in International Conference on Digital Telecommunications, 2010.

[43] R. Noldus, and P. Bukhave "Experience gained from IMS based Fixed Mobile Converged (FMC) network deployment, " in International Conference on Intelligence in Next Generation Networks, 2009.

[44] J. Cho, J.Hwang, and N. Kim "Novel FMC network composition methods in 3W (WiBro, WiFi and WCDMA) environments, " in International Conference on Advanced Communication Technology (ICACT), 2010.

[45] M. Choi, J. Ryu, and S. Jin, "Design and Implementation of Terminal System Based on the Next Generation FMC Using Wi-Fi and WiBro, " in International Conference on Control, Automation and Systems, 2010

[46] S. Gosselin, A. Pizzinat, X. Grall, D. Breuer, E. Bogenfeld, A. Hamidian, J. Torrijos Gijń, and N. Fonseca, "Fixed and Mobile Convergence: Which Role for Optical Networks?, " in Optical Fiber Communication Conference (OFC/NFOEC), 2015.

[47] R. Martinez, R. Vilalta, R. Casellas, and R. Muõz "Experimental Validation of a SDN Orchestrator for the Automatic Provisioning of Fixed and Mobile Services, " in European Conference on Optical Communications (ECOC), 2015.

[48] F. A.Lopes, M. Santos, R. Fidalgo, S. Fernandes, "A software engineering perspective on SDN programmability," IEEE Communications Surveys Tutorials, vol. 18, no. 2, pp. 1255 –1272, 2nd Quart. 2016.

[49] H. Farhady, H. Lee, and A. Nakao, "Software-defined networking: A survey," Computer Networks, vol. 81, pp. 79 –95, Apr. 2015.

[50] D. Kreutz et al., "Software-defined networking: A comprehensive survey," Proc. IEEE, vol. 103, no. 1, pp. 14-76, Jan. 2015.

[51] F. Hu, Q. Hao, and K. Bao, "A survey on software-defined network and OpenFlow: From concept to implementation," IEEE Communications Surveys Tutorials, Vol. 16, No. 4, pp. 2181-2206, 4th Quarter 2014.

[52] L. Cui, F. R. Yu, and Q. Yan, "When big data meets software-defined networking: SDN for big data and big data for SDN," IEEE Netw., vol. 30, no. 1, pp. 58-65, Jan./Feb. 2016.

[53] J. A. Wickboldt, W. P. De Jesus, P. H. Isolani, C. B. Both, J. Rochol, L. Z. Granville, "Software-defined networking: Management requirements and challenges," IEEE Commun. Mag., vol. 53, no. 1, pp. 278-285, Jan. 2015.

[54] A. S. Thyagaturu; A. Mercian, M. P. McGarry, M. Reisslein, W. Kellerer: "Software Defined Optical Networks (SDONs): A Comprehensive Survey", IEEE COMMUNICATIONS SURVEYS & TUTORIALS, vol. 18, no. 4, pp. 2738-2786 FOURTH QUARTER 2016.

[55] P. Bhaumik et al., "Software-defined optical networks (SDONs): A survey," Photon. Netw. Commun., vol. 28, no. 1, pp. 4-18, 2014.

[56] S. Gringeri, N. Bitar, and T. J. Xia, "Extending software defined network principles to include optical transport", IEEE Communications Magazine, vol. 51, no. 3, pp. 32-40, 2013

[57] T. P. de Souza, J. R. A. Cavalcante, A. Patel, M. E. Monteiro, J. Celestino Jr. "SONA: Software Defined Optical Networks Slicing Architecture", IEEE 31st International Conference on Advanced Information Networking and Applications, 2017

[58] M. Siqueira, J. Oliveira, G. Curiel, A. Hirata, F. Hooft, M. Nascimento, and C. Esteve Rothenberg, "An optical sdn controller for transport network virtualization and autonomic operation," in IEEE Globecom Workshops, pp. 1198-1203, 2013.

[59] S. Das, G. Parulkar, N. McKeown, P. Singh, D. Getachew, and L. Ong, "Packet and circuit network convergence with OpenFlow," in Optical Fiber Communication Conference/ Nat. Fiber Optic Engineers Conf. (OFC/NFOEC), 2010, paper OTuG1.

[60] L. Liu, D. Zhang, T. Tsuritani, R. Vilalta, R. Casellas, L. Hong, I. Morita, H. Guo, J. Wu, R. Martínez, and R. Muñoz, "First field trial of

an OpenFlow-based unified control plane for multilayer multigranularity optical switching networks", in Optical Fiber Communication Conf. and Expo. and the Nat. Fiber Optic Engineers Conf. (OFC/NFOEC), 2012, paper PDP5D.2.

[61] Nabil Bitar, "Software Defined Networking and applicability to access networks", in Optical Fiber Communication Conference (OFC), 2014

[62] K. Kerpez and G. Ginis "Software-Defined Access Network (SDAN)" in 48th Annual Conference on Information Sciences and Systems (CISS), 2014

[63] H. Yang, Y. Zhao, J. Zhang, R. Gu, J. Wu, J. Han, Y. Yao, Y. Lin, Y. Lee, and Y. Ji, "Experimental demonstration of remote unified control for OpenFlow-based software defined access optical networks" in European Conference on Optical Communications (ECOC), proceedings, 2013

[64] K. Kondepu, A. Sgambelluri, L. Valcarenghi, F. Cugini, P. Castoldi, "An SDN-based Integration of Green TWDM-PONs and Metro Networks Preserving End-to-End Delay", in The Optical Fiber Communication Conference and Exhibition (OFC), 2015

[65] D. Ren, S. Wu, L. Zhang, "A New Wavelength Optimization and Energy-Saving Scheme Based on Network Coding in Software-Defined WDM-PON Networks", in Journal of Optical Communications, Volume 37, Issue 3, pp. pp.265-269, 2016

[66] A. B. Sassi, M. S. Gavidia, E. L. Fernandes and M. R. Nascimento, "Integrated management of 10G-PON network element using NETCONF and OpenFlow," 10th International Conference on Network and Service Management (CNSM) and Workshop, 2014.

[67] ITU-T Recommendation Y.1541, International Communication Union ITU-T, 2011

[68] A. Rostami, P. Ohlén, M. S. Santos and A. Vidal, "Multi-Domain Orchestration across RAN and Transport for 5G," in Proceedings of theACM SIGCOMM Conference, pp. 613-614, 2016.

[69] M. Fiorani, Ahmad Rostami, L. Wosinska, P. Monti, "Transport Abstraction Models for an SDN-Controlled Centralized RAN," in IEEE Communications Letters, Volume 19, Issue 8, pp. 1406 - 1409, 2015

[70] H. Lan, H. Gu; X. Yu, X. Wang, "OROF: A hybrid topology with high scalability for data center networks," in International Conference on Optical Communications and Networks (ICOCN), 2015

[71] G. L. Vassoler, M. H.M. Paiva, M. R.N. Ribeiro, M. E.V. Segatto, "Twin Datacenter Interconnection Topology," in IEEE Micro, Volume 34, Issue 5, 2014

[72] W. Renqun, P. Li, "Hyper-DC: A rearrangeable non-blocking data center networks topology," in IEEE International Conference on Communication Software and Networks (ICCSN), 2016

[73] N. Mirfakhraei, Y. Tang, "Performance analysis of Benes networks under nonuniform traffic," IEEE International Conference on Communications, 1996

[74] A. Thamarakuzhi; J. A. Chandy, "Design and Implementation of a Nonblocking 2-Dilated Flattened Butterfly Switching Network," in IEEE Latin America Transactions, Volume 9, Issue 4, pp. 557 - 564, 2011

[75] T. Szymanski, ""Hypermeshes" : Optical Interconnection Networks for Parallel Computing," in Journal Of Parallel and Distributed Computing, Vol. 26, pp. 1-23, 1995

[76] C. M. Caba, J. Soler, L. Ditman "Software Defined Networking: Applicability and Service Possibilities," PhD Thesis, Technical University of Denmark, 2016

[77] M. Dallaglio, N. Sambo, F. Cugini, P. Castoldi, "Control and Management of Transponders With NETCONF and YANG," in Journal of Optical Communications and Networking, Vol. 9, N0. 3, March 2017.

[78] ORCHESTRA project, European Community's Horizon Program 2020,under grant agreement 645360, `http://www.orchestraproject.eu`.

[79] V. Lopez, O. Gerstel, R. Casellas, A. Farrel, D. King, S, López-Buedo, A. Cimmino, R. Morro, J. Fernandez-Palacios "Adaptive Network Manager: Coordinating Operations in Flex-Grid Networks," in International Conference on Transparent Optical Networks (ICTON), 2014.

[80] IDEALIST project, European Community's Seventh Framework Program FP7-2007-2013 through the Integrated Project (IP) IDEALIST under grant agreement 317999, http://www.ict-idealist.eu.

[81] L. Liu, R. Muñoz, R. Casellas, T. Tsuritani, R. Martínez, I. Morita, "Openslice: an OpenFlow-based control plane for spectrum sliced elastic optical path networks," Optics Express 21, 4194-4204, 2013

[82] A. Giorgetti, F. Paolucci, F. Cugini, and P. Castoldi "Dynamic Restoration With GMPLS and SDN Control Plane in Elastic Optical Networks" Journal of Optical Communications and Networking, Vol. 7, Issue 2, 2015

[83] J. Yin, J. Guo, B. Kong, and Z. Zhu, "Demonstration of survivable vSD-EON slicing with automatic data plane restoration to support reliable video streaming," in The Optical Fiber Communication Conference and Exhibition (OFC) pp. 1-3, 2017.

[84] A. Asensio, L. Gifre, M. Ruiz, L. Velasco "Carrier SDN for Flexgrid-Based Inter-Datacenter Connectivity" in International Conference on Transparent Optical Networks (ICTON), 2014

[85] Y. Yoshida, et al., "First Demonstration of Cognitive SDN Orchestration: A Real-time Congestion-aware Services Provisioning over OFDM-based 400G OPS and Flexi-WDM OCS Networks " (PDP) in The Optical Fiber Communication Conference and Exhibition (OFC), 2016

[86] M. Garrich, A. Bravalheri, M. Magalhães, M. Svolenski, X. Wang, Y. Fei, A. Fumagalli, D. Careglio, J. Solé-Pareta, J. Oliveira "Demonstration of Dynamic Traffic Allocation in an SDN-enabled Metropolitan Optical Network Test-bed" in International Conference on Optical Network Design and Modeling (ONDM), May, 2016.

[87] M. Channegowda, R. Nejabati, and D. Simeonidou "Software-Defined Optical Networks Technology and Infrastructure: Enabling Software-Defined Optical Network Operations," in Journal of Optical Communications and Networking, Vol. 5, Issue 10, pp. A274-282, 2013

[88] D. King, A. Farrel, "A PCE-based architecture for application-based network operations," IETF RFC 7491, 2015.

[89] O. Gonzalez de Dios, R. Casellas, F. Zhang, X. Fu, D. Ceccarelli and I. Hussain, "Framework and Requirements for GMPLS-Based Control of Flexi-Grid Dense Wavelength Division Multiplexing (DWDM) Networks," IETF Request for Comments: 7698, November 2015.

[90] J. E. Vergara, Daniel Perdices, V. Lopez, O. Gonzalez de Dios, D. King, Y. Lee, G. Galimberti, "YANG data model for Flexi-Grid Optical Networks", IETF draft-vergara-flexigrid-yang-02, Internet Draft, March. 2017.