A programmable Ethernet transport *Packetponder* using common compact form factor pluggable tunable transceivers to support novel DWDM architectures

Julie Raulin[1,2,*], Gawen Davey[3], Yuliya Verbishchuk[1,2,†], Alexander Jeffries[3], Cormac J. Sreenan[2] and Fatima C. Garcia Gunning[1]

[1]Photonics Systems Group, Tyndall National Institute, University College Cork, Cork T12 R5CP, Ireland.
[2]School of Computer Science and Information Technology, University College Cork, Cork T12 XF62, Ireland.
[3]APS Networks B.V., Wilhelmina van Pruisenweg 104, 2595 AN Den Haag, The Netherlands.
[†] now with VMWare, Cork, Ireland.
[*] julie.raulin@tyndall.ie

Abstract: We introduce a *packetponder* comprising a programmable packet switch with P4 ASIC containing a mixture of "grey" and tunable DWDM pluggable transceivers that, combined with ROADMs, introduces novel possibilities for Ethernet transport architectures. © 2023 The Author(s)

1. Introduction

There are three recent innovations that are set to transform optical networking. These are: 1) Hardware disaggregation, open source software and open application programming interfaces (APIs) - that usefully expose previously opaque mechanisms within whitebox network devices; 2) Compact, coherent optical pluggable transceivers - that are attracting interest from several telecommunications operators [1]; and 3) Programming Protocol-independent Packet Processors (P4) - an open source, domain-specific programming language - that specifies how data plane devices (switches, routers, FPGAs, etc.) can process and forward packets [2]; and can be combined collectively in novel ways.

In this paper, we introduce and demonstrate, for the first time to our knowledge, the concept of *packetponder*, where these three disparate innovations can be collectively harnessed to combine programmable Ethernet transport cross-connect supporting wavelength conversion and 3R optical-electrical-optical (OEO) regeneration within a single network element. The *packetponder* can be usefully complemented by DWDM ROADMs for analog optical wavelength switching and amplification.

2. Background and Concept

Optical transport transitioned from SONET/SDH to optical transport network (OTN) a.k.a. "digital wrappers" to support an array of cell- or frame-based L2 clients that included ATM, Frame Relay and Ethernet [3]. Ethernet dominated commercially and eventually prevailed, and so OTN cross-connects (XCs) used generic framing procedure (GFP) [4] to encapsulate/multiplex ethernet client frames in linecards surrounding a fixed-cell "universal switching" fabric [5]. An Ethernet transponder could be considered as an OTN XC variant that cross-connects between the client-side and a line-side. The line-side normally combines framing, DSP, and analog optics with wavelength tunability.

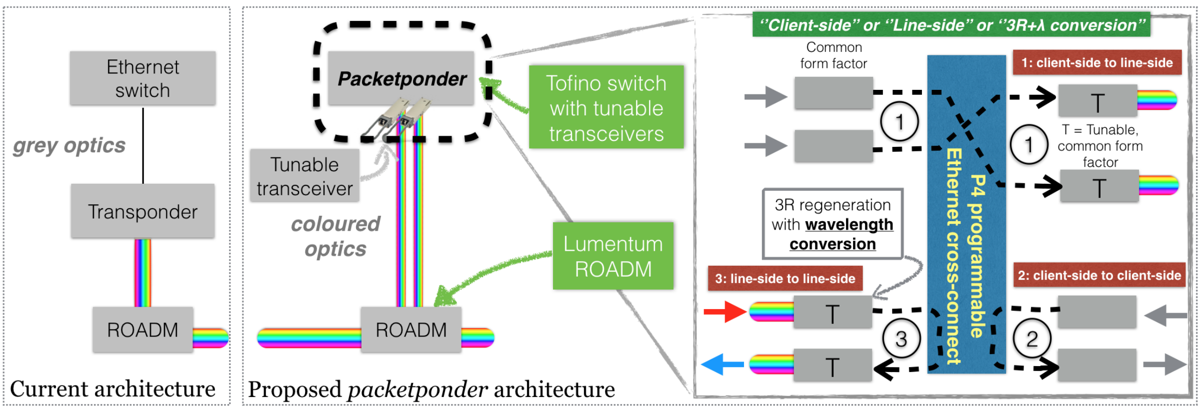


Fig. 1: Ethernet transport evolution: (a) current architecture with use of transponders; (b) proposed *packetponder* architecture; and (c) *packetponder*'s functionality: ① client-side to line-side - a transponder; ② client-side to client-side - Ethernet cross-connect; ③ line-side to line-side - a P4 programmable opaque optical cross-connect + 3R OEO regeneration with wavelength conversion.

The hyperscalers, including Google and Amazon, consume top-of-rack, aggregation Ethernet switches of increasing bandwidth capacities and decreasing power consumption per Gbit/s within their warehouse-scale datacentres. They collectively underwrite the continuous development of monolithic, merchant-silicon Ethernet ASICs of monotonically increasing bi-section bandwidths of several tens of Tbit/s that are contained in 1RU/2RU "Pizzaboxes". Carrier-grade, connection-oriented Provider Backbone Bridge - Traffic Engineering (PBB-TE, IEEE 802.1Qay-2009) is a mature and innovative technology. It provides a distributed Ethernet cross-connect that uses "mac-in-mac" (Ethernet-over-Ethernet) to encapsulate/multiplex Ethernet client frames using monolithic, merchant-silicon Ethernet ASICs. Forwarding is determined by a 4-tuple of header fields. P4 further enhances the possibilities of carrier-grade, connection oriented PBB-TE by enabling forwarding based on a larger superset of user-defined header fields. It also has an open, intuitive programming API for great flexibility.

Meanwhile, the commercial availability of tunable pluggable OpenZR+ transceivers that provide 100G–400G single-wavelength optical ports dedicated to the transport of Ethernet clients over DWDM networks is transformational. Despite challenges that currently preclude the incorporation of OpenZR+ in Ethernet switches, with the 14.5W power ceiling for QSFP-DD port cages in Ethernet switches being the most notable impediment to adoption [1], we anticipate Silicon CMOS DSP ASICs that exploit `5nm` and `3nm` nodes using extreme ultraviolet (EUV) lithography will reduce the power consumption of standard form-factor (QSFP-DD or OSFP) transceivers [6].

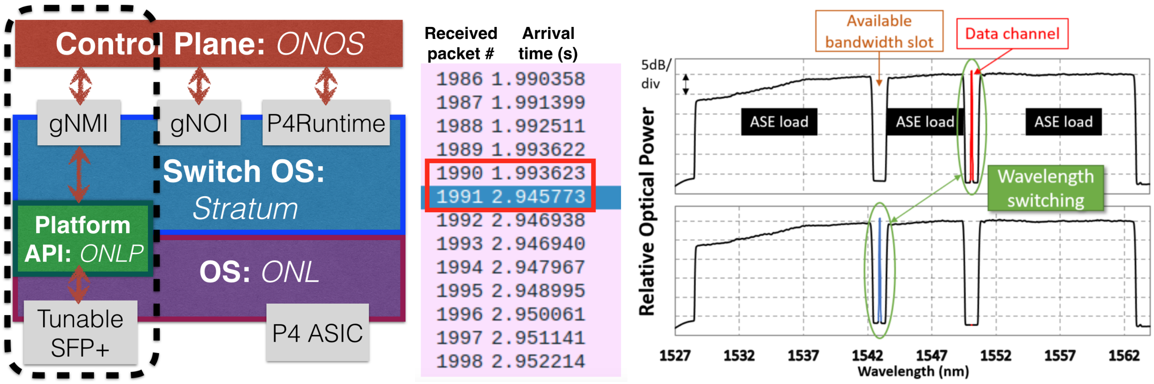
The development of commercial switches that combine P4 programmable switching ASICs with tunable, pluggable OpenZR+ transceivers within standard form factor frontplate cages is inevitable after power constraints are addressed. It obviates the requirement for a separate interposing transponder between the Ethernet switch and the ROADM, shown below in Fig. 1(a), by absorbing its function (Fig. 1(b)). The proposed *Packetponder* shown in Fig 1(c) rationalises the network architecture and supports existing and new functionalities: ① client-side to line-side; ② client-side to client-side; ③ line-side to line-side. The latter introduces wavelength conversion with 3R OEO regeneration that can be periodically invoked to extend the end-to-end reach beyond the analog transmission limit and/or relax the requirement for end-to-end wavelength continuity. It also removes the need for an extensive inventory of fixed wavelength transceivers. This flexible and intelligent provisioning can be achieved by tightly merging L0 capabilities to the L2/L3 management plane [7,8]. In this work, we manage the *packetponder* and the pluggables using open APIs, and we demonstrate a full stack management plane to enable wavelength conversion, and also OEO regeneration.

3. Experimental demonstration

Our proof-of-concept demonstrator used: an APS-Networks (BF6064X-T) switch that contained a P4 Tofino ASIC; a Lumentum OpenROADM (ROADM-20); and commercial Lumentum 10G SFP+ tunable transceivers (gray and tunable) *in lieu* of OpenZR+ pluggable optical transceivers that were not available to us. In order to emulate a fully loaded DWDM system, we used a filtered ASE source (or ASE load), emulating 80×50GHz grid spaced data channels, removing $4\times$ bandwidth slots (= 2×100GHz slots) to emulate free available bandwidth slots, and hence demonstrate in-situ wavelength switching. Our aim was to use open management software and Fig. 2(a) outlines the programmable SDN stack used for the experimental validation. ONOS was selected as the control-plane software, with gNMI as the cotrol data streaming prototol and the Yet Another Next Generation (YANG) data modelling language. Stratum was chosen as the Network Operating System (NOS) for the switch, combining Open Network Linux (ONL) with ONL Platform (ONLP). We used ONOS to send a gNMI request to Stratum, which pushed the updated SFF-compliant configuration to the tunable transceiver using `PushChassisConfig` to change to the new frequency (or wavelength). Our choice of ONOS, Stratum and ONL/ONLP enabled full stack programmability of the APS switch. Alternative open, accessible SDN stack models could also be used (for example, SONiC).

In order to demonstrate the potential of the *packetponder* architecture, our experiment pushed *in-situ* a selected wavelength (or frequency) down the full stack, from controller to the pluggable tunable 10G SFP+ transceiver via the I^2C bus. This independently set/reset the wavelength of each individual line-side tunable 10G SFP+ transceiver - without the need for a transponder or back-to-back muxponder [9]. This wavelength switching is shown in Fig. 2(c), between available bandwidth slots across the optical spectrum with inherent 3R OEO regeneration. As the wavelength of the transceiver changes, the corresponding ROADM rules also need to be updated. This can be achieved through a central controller connected to both the switch and the ROADM, simultaneously allocating a new frequency (or wavelength) channel via gNMI and NETCONF respectively. Ethernet test packets were generated by Ostinato (`ostinato.org`), a software traffic generator emulator, sent to the ingress port of the P4 switch and transmitted from an egress port via a tunable transceiver at a rate of 1000 packets/s. The optical data was then sent through the ROADM and looped-back with a patchcord to the receiver port of the 10G Tunable SFP+ transceiver.

The packets were subsequently captured by Wireshark for measuring packet loss and the reconfiguration time during the experiment. At the initial state, the tunable SFP+ in the programmable switch was set to the frequency of



(a) (b) (c)

Fig. 2: Experimental settings and results: (a) SDN management stack; (b) Wireshark snapshot capture of received packets from Ostinato; (c) line-side spectrum at the initial state (top trace) and after reconfiguration (bottom trace), showing *in-situ* wavelength switching between dis-contiguous wavelength spectrum slots.

193,400GHz. While the system is running, similarly to a real network with packets transmitted through the optical path, we then run an ONOS script to simultaneously adjust the start and end frequency values of the ROADM and the tunable transceiver centre frequency. The SFP+ transceiver wavelength change was effected by the addition of a new command to the YANG model tree (`sfp-frequency`), which can be called by ONOS to enable the center channel frequency to be changed. The new center frequency was set to 194,300GHz (= 1542.9nm). Fig. 2(c) shows the line-side optical spectrum, showing the emulated DWDM channels (or ASE load), and wavelength switching from 1550.1nm to 1542.9nm. The wavelength switching process requires hardware changes both in the transceiver and the ROADM, temporarily interrupting traffic as expected. This was measured to be approximately 950ms, as shown in Fig. 2(b), likely caused by delays in the ROADM reconfiguration, in alignment with typically measured ROADM reconfiguraton times between 2 to 4s [10].

4. Conclusion

In this paper, we introduced a *Packetponder* concept with tunable common form factor pluggable transceivers in a programmable Ethernet switch. This arrangement supported a concurrent mixture of three programmable operational modes: 1) a *client-side* to *line-side* mapping mode of a traditional transponder; 2) a *client-side* to *client-side* mapping mode of a traditional Layer 2 Ethernet cross connect; and 3) a novel *line-side* to *line-side* mapping mode between tunable coherent optical ports that provided wavelength conversion with 3R OEO regeneration. This experimental proof-of-concept shows, for the first time, the potential of a *Packetponder* to rationalise and simplify optical network architectures. It provided novel wavelength conversion - to make the most effective use of bandwidth; 3R OEO regeneration - inherent in the P4 switch ASIC; and full stack Open SDN management.

The authors would like to acknowledge P. Srivats for support with the Ostinato packet generator; B. O'Connor, M. Pudelko and Y. Tseng for support with Stratum; and our IPIC colleagues for useful network discussions. This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant numbers: 13/CDA/2103, 12/RC/2276-P2 (SFI Centre IPIC), 13/RC/2077-P2 (SFI Centre CONNECT).

References

1. TIP MANTRA Whitepaper: "IPoWDM convergent SDN architecture - motivation, technical definition and challenges, v1.0" (2022).
2. P. Bosshart et al.: "P4: programming protocol-independent packet processors", ACM SIGCOMM Comput. Comm. Rev. 44, pp. 87-95 (2014).
3. M. N. Ellanti et al.: "Next Generation Transport Technologies", in Next Generation Transport Networks Data, Management, and Control Planes, Chapter 4, pp. 230-261. ISBN 0-387-24067-5
4. E. J. Hernandez-Valencia et al.: "The Generic Framing Procedure (GFP): An overview", IEEE Comms. 40(5) (2002).
5. C. Baidong et al.: "ODU path protection in a disaggregated OTN switching system", US Patent 10,469,168 B2 (2019).
6. "TSMC logic technology": `https://www.tsmc.com/english/dedicatedFoundry/technology/logic`
7. D. Scano, et al.: "Hierarchical Control of SONiC-based Packet-Optical Nodes encompassing Coherent Pluggable Modules", ECOC'21, paper Tu2F-Demo.3 (2021).
8. O. Gerstel, et al.: "A Control Hierarchy for Integrated Packet-Optical Networks Utilizing Pluggable Transceivers", OFC'22, paper M4F-2 (2022).
9. A. Gumaste, et al.: "Hardware Comparison of Xponders and ZR+ in Metro and Core Networks with Mixed packet and OTN Traffic", OFC'22, paper Th1F-2 (2022).
10. A. Giorgetti, et al.: "Control of open and disaggregated transport networks using the Open Network Operating System (ONOS)", JOCN 12 (2), pp. A171-A181 (2020).