Optimal deployments of 400 Gb/s multihaul CFP2-DCO transponders in transparent IPoWDM core networks

Thierry Zami[(1)], Bruno Lavigne[(1)]
(1) Nokia, Route de Villejust, 91620 Nozay, France
thierry.zami@asn.com

Abstract: By comparing different strategies for deploying pluggable CFP2-DCO transponders interconnecting 400 Gb/s ports of distant IP routers in core WDM networks, we identify the one minimizing the number of required IP ports and/or of transponders per Gb/s. © Nokia 2022
OCIS codes: (060.1155) All-optical networks; (060.6718) Switching, circuit; (060.4256) Networks, network optimization.

1. Introduction

400 Gb/s coherent pluggable Elastic Optical Transponders (EOT) are eagerly expected to notably decrease the cost per Gb/s in WDM networks thanks to lower footprint, power consumption and capital expenditure than usual all-in-one EOTs, and above all thanks to direct pluggability into IP routers [1]. Various strategies of deployments are envisioned for this cost reduction and management simplification, from opaque networks to fully transparent ones [2][3]. In between these 2 opposite cases, the mean ratio of optoelectronic regeneration (RG) per optical connection can be chosen to adjust the average length and spectral efficiency of the transparent sub-Light Paths (LP) delineated by the regenerators. In that context, we examine the optimal setting for deploying 400 Gb/s multi-haul coherent CFP2-DCO EOTs [4] in the G50 and IND71 core optical fiber networks based on IP over WDM technology (IPoWDM), depending on the set of possible channel data rates and on the rules for RG.

2. Study assumptions

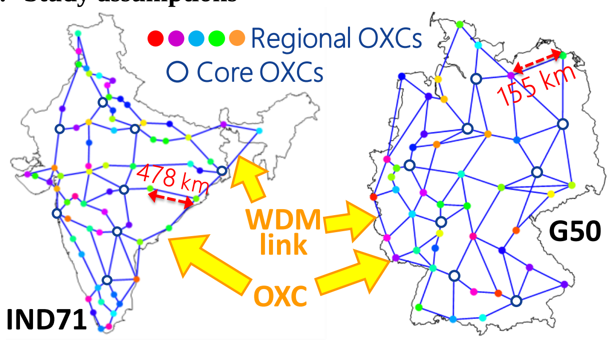

Figure 1: The G50 and IND71 network topologies

Figure 2: Example of a dual CFP2-DCO muxponder sled suitable as back-to-back WDM regenerator

The 400 Gb/s CFP2-DCO development has been initiated in the wake of the 400ZR OIF standard to accommodate optical connections across meshed WDM regional networks. Because of the distances its most emblematic 400 Gb/s channel can bridge, it evidently suits metro networks where most LPs are shorter than 500 km. Its relevance remains to be clarified for larger WDM core networks like G50 and IND71 shown in Fig. 1. G50 [4] (resp. IND71) is a German (resp. Indian) backbone WDM network having 50 (resp. 71) wavelength-routing Optical Cross-Connects (OXC) and 88 (resp. 97) links, each made of 2 counter-directional suites of fiber spans. These fibers are standard single mode fibers with 0.22 dB/km loss. Each total span loss is compensated by a dedicated erbium doped fiber amplifier featuring 5.5 dB noise figure. The mean/max span lengths are 45.6/75 km for G50 and 57.6/75 km for IND71. The electronically assisted coherent detection of the EOT offsets the optical dispersions linearly cumulated along the transmission LPs and performs forward error correction. Each fiber can have up to 64 WDM carriers 75 GHz-spaced over the 4800 GHz-wide C-band transmission window. The 4 carrier data rates supported by the simulated CFP2-DCO EOT are 100 Gb/s at 30 GBaud and 200, 300 or 400 Gb/s at 60 GBaud to transport GbE traffic from IP routers. The corresponding EOT performance is in-line with the specifications of [6]. Each OXC is based on the Route&Select layout [7] and associated with an IP router handling the added/dropped traffic with 400 Gb/s IP ports equipped with 400 Gb/s CFP2-DCO multi-haul EOT rather than QSFP-DD ones. Because CFP2-DCO EOTs transmit higher emitted power which better suits the OXCs in terrestrial core WDM networks. Although an IP router can exhibit a plethora of distinct port capacities, we assume IP ports handling up to 400 Gb/s. Because we envision such a medium-term widespread evolution for efficient cost reduction per Gb/s in the IP layer, even if most of the demands of Ethernet traffic remain 100 GbE ones. Internal optical amplification in OXC ensures at least -10 dBm channel received power for optimal

detection up to 400 Gb/s carrier data rate. The physical Quality of Transmission (QoT) of each transparent optical connection allocated by our network simulation accounts for the linear optical noises, the non-linear ones due to WDM transmission and the impact of all the optical filters traversed by the optical channel under test [8].
Hierarchical traffic is simulated in these 2 networks with 2 classes of OXC shown in Fig. 1 where each network has 8 core OXCs depicted as white empty points. These core OXCs exchanged the largest throughputs between themselves in a fully meshed virtual topology. The other OXCs are regional ones sketched as colorful points in Fig. 1. Each one only exchanges traffic with its 2 closest core OXCs. For each simulated network, the traffic matrix is randomly drawn as follows: connections are bidirectional and symmetrical; there is a fixed set of 400 Gb/s connections between each pair of core OXCs and 100 Gb/s connections between each regional OXC and its 2 closest core OXCs; the remaining part of the demanded traffic is random with 33% (resp. 66%) probability of core OXC to core OXC (resp. core OXC to regional OXC) connection carrying 100 to 400 Gb/s (resp. 100 Gb/s). Inside this 100-to-400 Gb/s interval, the random throughput is uniformly selected with 100 Gb/s granularity. The fixed part of traffic is served before the random part, for both one connection after the other one. We assume IP routers also do end-to-end traffic grooming and inverse multiplexing over several optical subcarriers when the demanded capacity exceeds the maximum capacity of a single optical carrier along its tested LP. Then, these subcarriers follow the same LP and are regenerated in the same OXCs. RG can overcome too poor channels QoT or wavelength contention along their LP. RG can also increase the spectral efficiency by shortening the lengths of transparent LPs. If a suitable LP with enough available spectral resource cannot be found to serve a request for connection, it is discarded. Each simulation runs until at least 1% rejection of the overall demanded throughput. For a pair of simulated network and traffic, the Maximum Network Capacity (MNC) is defined as the total throughput served up to this 1% of blocking. We also report the mean number of required CFP2-DCO EOTs (N_{EOT}) and 400 Gb/s IP ports (N_{IP}) at this 1% blocking situation, both per 400 Gb/s transmitted from its source to its destination OXC. N_{EOT} and N_{IP} account for the regenerating EOTs and they cannot be smaller than 2, since for a fully filled 400 Gb/s transparent connection, 2 pairs of IP ports+400 Gb/s EOTs are needed, one at the source node and one at the destination node. In case of one intermediate RG for a 400 Gb/s service, then N_{EOT}=4 and N_{IP}=4. So, the closer to 2 N_{EOT} and N_{IP} are, the better. MNC, N_{EOT} and N_{IP} metrics reported in the remainder are averaged over 100 distinct random traffic draws fulfilling the previously described distributions.

3. Simulation results and discussion

We simulate the 5 following network strategies: i) $\mathbf{S_{Opaque}}$ meaning that each channel goes back into the IP layer at each network node it crosses, enabling intermediate traffic regrooming; ii) $\mathbf{S_{400}}$ for G50 transparent network with only 400 Gb/s carriers keeping the ultimate IP port and EOT capacities fully aligned; iii) $\mathbf{S_{300}}$ for IND71 transparent network only with 300 Gb/s carriers (S_{400} is not applied to IND71 since its 27% of its fiber links are too long to be covered by 400 Gb/s CFP2-DCO carriers); iv) $\mathbf{S_{FLEX}}$: transparent network with channel data rate flexibility minimizing N_{EOT} & N_{IP}; v) $\mathbf{S'_{FLEX}}$: also transparent network but maximizing MNC. For each of the strategies S_{400}, S_{300}, S_{FLEX} and S'_{FLEX}, we investigate RG either through the IP routers, called "IP-RG", or with dedicated sleds as depicted in Fig. 2, called "no IP-RG". This latter setup saves IP ports when a channel is regenerated without changing its data rate. Since such a no IP-RG only goes through an electrical backplane without IP packet processing, it also reduces latency. However, unlike IP-RG, it cannot change the carrier data rate for better capacity breakdown between subcarriers from the 2 distinct sides of the RG point. IP-RG setup means one-to-one mapping between EOTs and 400 Gb/s IP ports and so it implies N_{EOT}=N_{IP}. For any of these strategies, each IP port can dial its throughput lower than 400 Gb/s to directly fit the maximum capacity enabled by the QoT throughout the LP of the WDM carrier it is connected to.
The results for our 5 strategies appear in Table 1. As expected, S_{Opaque} leads to the largest MNCs, but also yields very high N_{EOT} and N_{IP}. S_{400} strategy applied to G50 saves at least 42% of EOTs and IP ports per Gb/s (4.05 vs. 7) with nearly no impact on MNC. This is because S_{400} bypasses some of the useless S_{Opaque} regenerations. N_{IP} reduction even reaches 70% (2.06 vs. 7) under no IP-RG condition. This quantifies the superiority of WDM transparency against opacity for a backbone network like G50 once its traffic load is high enough thanks to at least end-to-end electrical

Table 1: *Network performance for G50 and IND71 topologies with 400 Gb/s CFP2-DCO interfaces in an IPoWDM context.*

	G50				**IND71**			
	$\mathbf{S_{Opaque}}$	$\mathbf{S_{400}}$	$\mathbf{S_{FLEX}}$	$\mathbf{S'_{FLEX}}$	$\mathbf{S_{Opaque}}$	$\mathbf{S_{300}}$	$\mathbf{S_{FLEX}}$	$\mathbf{S'_{FLEX}}$
MNC (Tb/s)	433.3	429.3	336.9	373.2	271.6	253.5	193.5	243.8
"IP-RG" N_{EOT}&N_{IP} per 400 Gb/s	7	4.05	2.69	3.2	9.76	4.88	3.9	4.9
"no IP-RG" N_{IP} per 400 Gb/s	N.A.	2.06	2.39	2.22	N.A.	2.82	3.2	2.75
Percentage of regenerated connections	72%	48.9%	9.85%	32.9%	81.3%	38.51%	14.22%	60.7%
Percentage of 100 Gb/s channels	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0.24%**	**0%**
Percentage of 200 Gb/s channels	**0%**	**0%**	**2.9%**	**1.1%**	**0%**	**0%**	**38.9%**	**0%**
Percentage of 300 Gb/s channels	**0%**	**0%**	**42.2%**	**16.5%**	**27.3%**	**100%**	**48%**	**81.9%**
Percentage of 400 Gb/s channels	**100%**	**100%**	**54.9%**	**82.4%**	**72.7%**	**0%**	**12.86%**	**18.1%**

traffic grooming. Applying S_{FLEX} on G50 can further reduce N_{EOT} and N_{IP} down to 2.69 under IP-RG condition, but it also decreases MNC by 21.5% (429.3 vs 336.9 Tb/s). Besides, S_{FLEX} does not save as many IP ports per Gb/s as S_{400} if "No IP-RG" approach is adopted, even if it considerably lowers the ratio of regenerated connections. This is because with S_{400}, all the regenerated channels carry 400 Gb/s and then their regeneration can be done by means of sleds like in Fig. 2. This uniformity of channel data rate even in a transparent context is allowed thanks to the small G50 diameter relatively to IND71. Whereas with S_{FLEX} the channel data rates are more distributed between 300 and 400 Gb/s due to the longer transparent LPs (see bottom lines in Table 1) to save regenerating EOTs. Hence, residual RG often happens between 300 and 400 Gb/s carriers and this is inevitably done by IP ports. S'_{FLEX} better trades off than S_{FLEX} for G50 as it mitigates the MNC decay from S_{400} and still notably improves N_{IP} from 2.39 down to 2.22. S_{400} or S'_{Flex} performances reported in Table 1 indicates multi-haul CFP2-DCO EOTs can be implemented in transparent optical networks like G50 more efficiently than an opaque version by grooming traffic on an end-to-end basis to sufficiently fill the optical channel throughputs.

IND71 simulations result in the same trends as for G50 but with different proportions. Hence, MNC still peaks with S_{Opaque} but with only 7% extra MNC than S_{300} while it needs 100% extra N_{EOT} and N_{IP} if regenerating via IP router, and even 246% more N_{IP} (9.76 vs 2.82) if regenerating with dedicated sleds. S_{FLEX} strategy also notably reduces N_{EOT} by at least 60% and N_{IP} by up to 67% compared to S_{Opaque}, at the cost of 28.7% MNC reduction (271.6 vs. 193.5 Tb/s). By fostering MNC instead of N_{EOT} reduction unlike S_{FLEX}, S'_{FLEX} only impacts MNC by 10.2% (271.6 vs. 243.8 Tb/s) while still reducing N_{EOT} and N_{IP} by 49.8% against S_{Opaque} under "IP-RG" rules. Because by more regenerating, S'_{FLEX} leads to shorter transparent LPs bridged in average by higher channel data rates, as quantified by the bottom lines of Table 1. This also raises probability of RG without changing channel data rate by means of sled instead of IP router. This explains the lower N_{IP} values under "No IP-RG" condition for S_{300} and S'_{FLEX}. Their networking performance for IND71 are close and both can be deemed as better than S_{FLEX} despite 26% higher needed N_{EOT} compensated by at least 26% better MNC (243.8 vs. 193.5 Tb/s) and 14% lower N_{IP} (2.75 vs. 3.2). Thus, when it comes to capital expenditure related to EOTs and IP ports, S'_{FLEX} remains more profitable for IND71 than S_{FLEX} if the cost of a CFP2-DCO pluggable interface is 26/14=1.85 times less expensive than the cost of a 400 Gb/s IP port. Despite the possibility to further allocate 400 Gb/s carriers, S'_{FLEX} does not reach higher MNC than S_{300}. This stems from suboptimal gradual end-to-end traffic grooming for the regenerated connections combining 300 Gb/s and 400 Gb/s carriers on disjoint parts along their LP. This slightly affects the network performance even with the higher-capacity 400 Gb/s carriers.

4. Conclusion

We have simulated planning of the G50 and IND71 WDM networks corresponding to 2 significantly different sizes, only equipped with 400 Gb/s IP routers and 400 Gb/s CFP2-DCO EOTs. Their optimization is mostly driven by the regeneration rule. If regeneration can be supported via two back-to-back 400 Gb/s CFP2-DCO modules, then the ratio of 400 Gb/s channels can be augmented in a transparent context without increasing the amount of IP ports. This network study illustrates that focusing only on reducing N_{EOT} might not be enough in IPoWDM framework. For the WDM core networks the diameter of which complies with typical 400 Gb/s transmission reach of pluggable CFP2-DCO multi-haul interfaces, N_{EOT} and N_{IP} could be straightforwardly co-optimized as illustrated by the G50 example. Because most connections will be served transparently, leading to a direct mapping between N_{EOT} and N_{IP}. But for larger networks, both N_{EOT} and N_{IP} should be handled individually for further network optimization, as done with S'_{FLEX} strategy. One possible effective exploitation of the CFP2-DCO elasticity could also consist in selecting a unique best channel modulation for a given network, depending on the QoT distribution of its LPs. Ultimately, bypassing IP routers for most regenerations will also be beneficial in terms of power consumption and overall service latency.

5. References

[1] S. Melle et al., "*Comparing IP-Optical Architectures & WDM Transport Technologies in Metro, Regional and Long-Haul Networks*", Paper M3E.2, OFC'2021, June 2021

[2] P. Wright et al., "*Cost Model Comparison of ZR/ZR+ Modules Against Traditional WDM Transponders for 400G IP/WDM Core Networks*", Paper Tu2H-2, ECOC'2020, December 2020

[3] T. Zami et al., "*Benchmarking of Opaque Versus Transparent Core WDM Networks Featuring 400ZR+ QSFP-DD or CFP2 Interfaces*", Paper We2K-7, ECOC'2020, December 2020

[4] B. Lavigne et al., "*20.8 Tb/s Transmission over 1200 km Using G654E Fibers, Hybrid Amplification and 400 Gb/s CFP2-DCO Interfaces*", Paper F4G.5, OFC'2021, June 2021

[5] G50 topology, http://sndlib.zib.de

[6] Open ZR+ MSA Technical Specification, Rev. n°1, September 2020

[7] B. Colling, "*New devices enabling software-defined optical networks*", IEEE Communications Magazine. Volume 51, Issue 3, Page 66–71 March 2013.

[8] T. Zami et al., "*Simple self-optimization of WDM networks based on probabilistic constellation shaping*", IEEE/OSA Journal of Optical Communications and Networking, Vol. 12, Issue 1, January 2020, https://doi.org/10.1364/JOCN.12.000A82