Role of Coherent System in the Next DCI Generation

D. Tauber
Lumentum Operations LLC., 1001 Ridder Park Dr, San Jose, CA 95131
dan.tauber@lumentum.com

Abstract: Coherent Transmission has been the standard for fiber optic transmission beyond 40 km for over a decade. We review its continuing role for DCI at 400 and 800 Gbps and higher rates.

1. Introduction and Network Architecture

As demand for global cloud services grows at a rapid rate, demand for high bandwidth interconnect between data centers also grows. The data center interconnect (DCI) optical architecture typically involves a single point to point span, with dense wavelength division multiplexed (DWDM) channels connected to each other in an amplified optical mux-demux architecture. For shorter links without multi-channel requirements, single wavelength links without amplifiers are also used. Longer reach DCI, in a multi-span point to point architecture, is also relevant, with rates between 100-400 Gbps in the current generation.

DCI data centers are typically separated by between 10 and 120 km, with most of the spans 80 km or less [1]. Three significant factors crucial for successful deployment of transceivers in DCI networks are (a) cost of optics (b) power consumption and (c) faceplate density. Due to the complexity of coherent technology with respect to power dissipation and size, DCI technology has traditionally required either separate transponder chassis that connect coherent transceivers to the pluggable client optics in routers or switches, or IM-DD transceivers that operate at lower rates and/or on multiple wavelengths - for example, 50 Gbps per wavelength/100 Gbps dual wavelength PAM-4 pluggable modules [2]. These solutions have been sufficient to support the cloud data internetworking requirements through the 2020 timeframe; however, the relentless growth in data demand, covering more and more 100GE clients, along with a ramp-up of 200GE and 400GE clients, has created urgency for an IP over DWDM infrastructure with high speed coherent transceivers integrated directly into routers and switches. Figure 1(a) shows the earlier generation of DCI, using coherent transponders connected via client optics to the router or switch. Figure 1(b) shows the new DCI architecture, where high data rate pluggable coherent transceivers eliminate the client optics and the coherent transponder-based chassis systems completely.

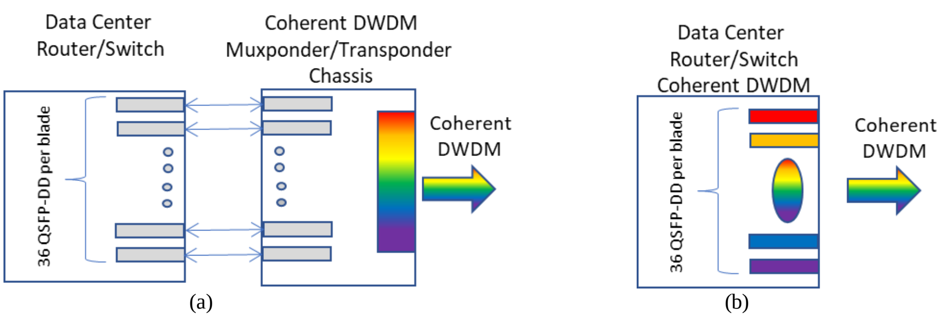

Figure 1. Transition of DCI from (a) separate coherent transponders chassis connected by IM-DD optics to router/switch to (b) IP over DWDM, where the coherent transceivers are integrated into the router/switch. Large cost, space, and power savings are achieved by collapsing the earlier generation into a single IP over DWDM platform.

2. Coherent Transceiver Requirements for DCI

2.1. Form Factor

The adoption of coherent transceivers in the DCI space comes with constraints on the transceiver design. Due to faceplate density requirements and the ubiquitous inclusion of QSFP-DD or OSFP plugs in routers and switches, coherent DCI transceivers at 400G and higher are being built in QSFP-DD and/or OSFP form factors. The 18.35 mm width/8.5 mm height of the QSFP-DD module (22.58mm/13.0mm for OSFP) is a design constraint that leads to innovations in package design and power dissipation reduction. Other form factors, such as CFP2-DCO are also used for DCI, depending on customer architecture.

To achieve small form factors with higher data rates, technology advances for minimizing DSP size are critical. CMOS nodes have moved from 7 nm for 400G to 5 nm for 800G, and it's expected that 1.6 Tbps and higher rates will utilize further reduction in node size to 3 nm. Compact packaging of optics and electronics is also critical, with tighter packaging approaches for optics and electronics for reduced size and improved thermal and RF properties.

2.2. Power Dissipation

Critically important is that these modules operate with low power consumption, especially for QSFP-DD, in the range of 20W, so that the switches can accommodate module cooling requirements, and to manage overall power consumption in the data center for environmental and cost reasons. Due to the power consumption challenges with coherent technology, partnerships between module suppliers and equipment vendors have realistically allowed for powers in the vicinity of 25W as rates increase [3].

As discussed in Section 2.1, coherent for next generation DCI uses ASICs with reduced CMOS DSP nodes sizes. Smaller node sizes are also necessary to minimize power dissipation. On the optics side, the transceiver must be designed with just enough transmitter optical power and OSNR for closing DCI links while also keeping power dissipation low. This leads to careful link engineering design, with TX output powers in the range of -10 dBm, rather than the 0 dBm range used in more traditional telecom applications, though powers above -10 dBm are also under consideration to provide more margin. The lower output power shortens the transmission reach, as the OSNR out of the booster amplifier is reduced with lower input, but with careful DCI link design compatible with the transceiver specifications, this is a manageable problem.

2.3. Interoperability

Another factor for facilitating adoption of coherent in DCI networks is line side multi-vendor interoperability, as is the norm in short reach intra-data center applications. The first demonstration of coherent DCI interoperability is the Optical Internetworking Forum (OIF) Implementation Agreement (IA) for the 400G ZR transceiver module [4], where multiple vendors agreed to use C-FEC forward error correcting code, transmitted on a 400G 16QAM constellation, with interoperable framing, and a client side that accepts 400GE clients; similar activities are underway in the OIF for 800G, and are anticipated for 1.6 and 3.2 Tbps also in the next few years. A variety of Ethernet clients will be supported including 100GE, 200GE, and 400GE, and an 800G client also. The Open ZR+ MSA [5] published a specification for extended reach 100-400G rates and continues to work on a variant for 400G on 75 GHz grid spacing. In parallel with this the IEEE project P802.3cw is also developing an amendment for 400 Gbps Ethernet over DWDM systems.

The OIF 400G ZR IA does not specify form factor. Modules other than QSFP-DD and OSFP can also be used for ZR applications if customer systems support different form factors. An example of a 400G ZR DP-16QAM constellation from a Lumentum CFP2-DCO module is shown in Figure 2.

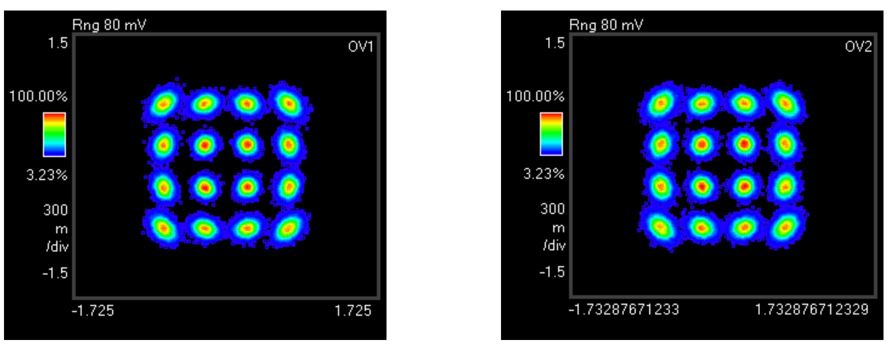

Figure 2. 400G ZR 16QAM X and Y constellation from Lumentum 400G CFP2-DCO.

400G ZR DCI deployments have ramped in 2021 at the major cloud companies, showing the need for coherent transceivers in the DCI space. Multiple reports have been released demonstrating vendor interoperability [6,7,8].

3. Transition from 400Gbps to 800Gbps and Higher Rates

Key optical parameters for both 400G and 800G DCI applications are shown in Table 1, covering optical transmit power, required OSNR at the receiver for amplified links, and minimum receive optical input power for DWDM amplified links with mux/demux. The chromatic dispersion that must be supported with minimal OSNR penalty is also shown. The 400G specs in the Table have been standardized by the OIF, while the 800G specs are in the process of standardization. At 1.6 Tbps one option is an integrated dual wavelength solution also operating at 120 Gbaud/16QAM; for this solution the optical specifications will be similar to 800G ZR specifications, albeit with dual carriers. 3.2 Tbps is further out, but could potentially leverage further integration for multiple carriers, and/or technology advances supporting 1.6 Tbps per wavelength in a dual wavelength solution.

Not shown in Table 1 are extended reach specifications for DCI. As is the case with the OpenZR+ for 400G, next generation is expected to support a longer reach multi-span 800G 16QAM mode, with reaches between 300-500 km, and a long distance 400G QPSK mode supporting 1000-2000 km.

Table 1. 400G ZR/800G ZR key optical specifications

	400G ZR	*800G ZR (draft)*
Modulation Format	16QAM	16QAM
Baud (Gbaud)	60	*120*
FEC Type	C-FEC	*OFEC or CFEC*
TX output power (dBm)	-10	*-9 to -13*
ROSNR (dB)	26	*26-29*
Rx input power (dBm) (amplified)	-12	*-12*
Chromatic Dispersion (ps/nm)	2400	*2400*

4. Next generation coherent module design considerations for DCI

In transitioning from 400G transmission at 60 Gbaud to 800G transmission per wavelength at 120 -130 Gbaud, and higher, we must consider key design factors for building a usable module. These parameters and associated design considerations are listed in the bullets below.

- Optical TX power: high enough to close DCI links while complying with size/power dissipation constraints
- Analog bandwidth of TX/RX chain: 50-70 GHz composite bandwidth (120 Gbaud), peaking in RF ICs, DSP equalization to optimize performance. For very high bandwidth (>128 Gbaud) InP and/or other novel optical materials
- RF interconnect between DSP, ICs, optics: compact packaging minimizing wire bond length, parasitics and reflections
- Modulator insertion loss & Vpi: Solutions with the compactness/packaging advantages of SiPh and the material benefits of InP
- DSP die/package size: purposeful design for small module integration, not constrained by traditional BGA/PCB based design
- Optics size/geometry: purposeful design for small module integration, considerations on height and dimensional alignment for optimizing IC and DSP connectivity
- Thermal Management: Packaging/heat sinking designs that minimize rise of DSP junction temperature relative to case, optimized TEC designs for temperature sensitive optics

The problem being solved in next generation DCI at 800Gbps and higher with coherent transmission is fundamentally an integration problem to enable DCI performance in form factors that support high faceplate densities compatible with next gen switch bandwidths of 25.6, 51.2, and 102.4 Tbps [9]. Having optical specs that ensure sufficient DCI link margin is critical for enabling robust network operation in the small form factor modules. Power dissipation must be managed with transceiver designers and manufacturers working closely with equipment designers and manufacturers to specify and meet budgets. Success will involve a thorough and comprehensive understanding of how state of the art laser, modulator and receiver technology works together with analog integrated circuits (ICs) and coherent DSPs. Package design of the optical and electrical components must be optimized to meet size and power requirements, to ensure success in a cost sensitive market.

5. References

[1] M. Filer, et al., J. Opt. Commun. Netw., **11** (10), C94-C108 (2019)

[2] R. Nagarajan, et al., J. Opt. Commun. Netw., **10** (7), B25-B36 (2018)

[3] QSFP-DD, Optimizing QSFP-DD Systems to Achieve at Least 25 Watt Thermal Port Performance (http://www.qsfp-dd.com/wp-content/uploads/2021/01/2021-QSFP-DD-MSA-Thermal-Whitepaper-Final.pdf) (2021)

[4] OIF, "Implementation Agreement 400ZR," https://www.oiforum.com/wp-content/uploads/OIF-400ZR-01.0_reduced2.pdf.

[5] OpenZR+ MSA, http://openzrplus.org/documents/

[6] https://www.arista.com/en/company/news/press-release/12161-pr-20210217

[7] https://www.lightwaveonline.com/optical-tech/transmission/article/14186222/acacia-communications-inphi-demo-400zr-interop-via-different-dsps

[8] https://www.fibre-systems.com/news/inphi-neophotonics-demo-400zr-interoperability-over-120km

[9] A. Srivastava, OFC '21, M5G.2, (2021)