Experience Track

Lightwave Fabrics: At-Scale Optical Circuit Switching for Datacenter and Machine Learning Systems

Hong Liu, Ryohei Urata, Kevin Yasumura, Xiang Zhou, Roy Bannon, Jill Berger, Pedram Dashti, Norm Jouppi, Cedric Lam, Sheng Li, Erji Mao, Daniel Nelson, George Papen, Mukarram Tariq, Amin Vahdat
Google
lightwave-fabrics@google.com

ABSTRACT

We describe our experience developing what we believe to be the world's first large-scale production deployments of lightwave fabrics used for both datacenter networking and machine-learning (ML) applications. Using optical circuit switches (OCSes) and optical transceivers developed in-house, we employ hardware and software codesign to integrate the fabrics into our network and computing infrastructure. Key to our design is a high degree of multiplexing enabled by new kinds of wavelength-division-multiplexing (WDM) and optical circulators that support high-bandwidth bidirectional traffic on a single strand of optical fiber. The development of the requisite OCS and optical transceiver technologies leads to a synchronous lightwave fabric that is reconfigurable, low latency, rate agnostic, and highly available. These fabrics have provided substantial benefits for long-lived traffic patterns in our datacenter networks and predictable traffic patterns in tightly-coupled machine learning clusters. We report results for a large-scale ML superpod with 4096 tensor processing unit (TPU) V4 chips that has more than one ExaFLOP of computing power. For this use case, the deployment of a lightwave fabric provides up to 3× better system availability and model-dependent performance improvements of up to 3.3× compared to a static fabric, despite constituting less than 6% of the total system cost.

CCS CONCEPTS

• **Networks** → **Network design principles**; • **Hardware** → **Networking hardware**;

KEYWORDS

Data center networks, Optical circuit switches, Machine learning

ACM Reference Format:
Hong Liu, Ryohei Urata, Kevin Yasumura, Xiang Zhou, Roy Bannon, Jill Berger, Pedram Dashti, Norm Jouppi, Cedric Lam, Sheng Li, Erji Mao, Daniel Nelson, George Papen, Mukarram Tariq, Amin Vahdat. 2023. Lightwave Fabrics: At-Scale Optical Circuit Switching for Datacenter and Machine Learning Systems. In *ACM SIGCOMM 2023 Conference (ACM SIGCOMM '23), September 10–14, 2023, New York, NY, USA.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3603269.3604836

ACM SIGCOMM '23, September 10, 2023, New York, NY, USA

ACM ISBN 979-8-4007-0236-5/23/09…$15.00
https://doi.org/10.1145/3603269.3604836

1 INTRODUCTION

Datacenter and campus-scale networks have been a key enabler for many advancements in large-scale computing over the last 20 years. Many ground-breaking infrastructure services ranging from distributed storage systems [17], programming models for large-scale data processing [11, 14], and machine learning (ML) [1] are at their core enabled by a datacenter network that provides low latency and abundant bandwidth among a large number of compute and storage servers. Similar networking capabilities at campus, regional and global scale have enabled increased-availability storage services and deployment archetypes [6, 13] as well as larger computing pools with superior efficiency and scaling properties.

During this period, the datacenter and regional network interconnects have been built around *electrical packet switching* (EPS), much like the Internet and local area networks. While packets may be transported optically, it is done along point-to-point paths over a fixed network topology. The physical topology changes rarely, primarily in response to network augments or upgrades. Since these networks serve applications with a wide range of performance requirements, they demand topologies that support a wide variety of traffic patterns. A common approach has been to rely on variants of Clos topologies [3, 24, 53] at datacenter and campus scale as shown schematically in Fig. 1a). While incredibly flexible, these networks are designed to admit worst-case permutation traffic patterns, and incur substantial cost, power, and latency to do so at scale. Fortunately, the journey to large-scale datacenters has included network control and cluster resource scheduling systems that have a central view of the applications and their communication needs [24].

A newer journey in supporting large-scale machine learning applications is showing substantial performance benefits (greater than 3× for some workloads) by shaping the configuration of the compute topology to match the communication patterns of the model (cf. §4.2.1).

For both of these pervasive use cases, we observe opportunities to co-optimize the resource allocation and network topology for the real-time communication patterns and achieve much higher performance without the substantial costs associated with the traditional

EPS-based network or for the case of ML, of using a static topology. However, these optimizations are only possible if we move away from the notion of a fixed topology as a cornerstone of networking. *In this paper, we share our design and production experience with what we believe to be the first at-scale deployment of adaptive optical circuit switching in both datacenter networks and ML clusters that can dynamically adjust topologies in response to real-time workload communication patterns.*

We call these reconfigurable networks **lightwave fabrics**, which are constructed using large-radix optical circuit switches and custom optical transceivers. Our large-scale deployment and substantiated benefits are contrary to conventional wisdom that lightwave fabrics cannot be built to the requisite levels of efficiency and reliability when considering the hardware challenges facing transceiver and OCS design, and the system challenges of deployment, availability, and cost efficiency.

Our production lightwave fabrics drive substantial gains for three emerging use cases: (i) datacenter networks that must support flexible traffic matrices for various workloads as well as interoperability among multiple generations of network technologies [47], (ii) ML supercomputers (called superpods) that want to leverage the simplicity, cost, and latency benefits of specialized topologies [25], and (iii) campus networks that must support a range of cluster-to-cluster communication patterns, shifting with the turnup and turndown of services, often spanning multiple clusters at today's scale. These use cases come together to form a hierarchical hybrid electrical/optical switching interconnect datacenter that benefits many different applications; in this paper we quantify the benefits in the context of large-scale ML training.

The rest of the paper is organized as follows. In §2, we motivate the key requirements for lightwave fabrics by describing how we use lightwave fabrics for scale-up and scale-out applications. §3 presents the design for the optical circuit switch (OCS) and optical transceivers. We evaluate the performance of hardware in §4.1 and the benefits of the lightwave fabric in §4.2, particularly focusing on ML. We present related work in §5, future work and other potential use cases for lightwave fabrics in §6, and conclude in §7.

This work does not raise any ethical issues.

2 THE OPTICAL DATACENTER

Our datacenters use lightwave fabrics for both general-purpose network connectivity and high- performance networks designed to support ML.

2.1 Datacenter Networks

The benefits of using a lightwave fabric to co-optimize the resource allocation and network topology for a datacenter network (DCN) have been well documented [8, 16, 31, 39, 40, 54, 56, 61, 62]. These benefits include being data-rate and wavelength agnostic, low latency, and extremely energy efficient.

Fig. 1 shows the evolution of a "spine-full" electrically packet switched (EPS) network to a "spine-free" lightwave network. Fig. 1a) illustrates a traditional datacenter network, with spine blocks (SPs) connecting homogeneous aggregation blocks (ABs). Fig. 1b) illustrates an evolved datacenter network incorporating a directly-connected lightwave fabric [47]. Besides 30% reduction in capex and 41% reduction in power, the lightwave fabric provides several other key benefits:

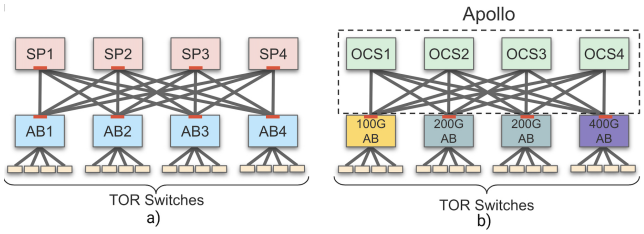

Figure 1: Evolution of Datacenter Networks. a) Traditional hierarchical datacenter network using spine blocks to connect aggregation blocks. b) Evolved heterogeneous datacenter network with spine blocks replaced by OCSes.

Topology Engineering: In this spine-free architecture, application-specific network topologies can be constructed via direct optical connections in addition to traditional switch-level traffic engineering. For example, reconfiguration of the fabric connectivity allows the optimization of inter-AB bandwidth when there is an increase in long-lived traffic demand between a particular set of ABs.

Fabric Expansion: With expansion capability, the size of the network fabric can be augmented incrementally, allowing an initial number of ABs to be deployed with additional ABs added to the fabric as needed ("pay as you grow").

Fabric Isolation: OCSes can create dynamic circuits to reconfigure the network topology and isolate performance across jobs and/or customers.

Rapid Technology Refresh: The expansion capability leads to the ability to connect different-generation ABs running at different data rates (and employing different EPSes and transceivers) to the same OCS. Interoperability between heterogeneous ABs is ensured through the compatibility of optical transceiver specifications across multiple generations of data rates leading to faster introduction of new technology.

2.2 Machine Learning Networks

To support the synchronous communication patterns and parameterizations used in large-scale ML models [1], the interconnection fabric can be scaled in two complementary ways. The first is to use an inter-chip-interconnection (ICI) fabric [27] to scale up tightly-coupled compute nodes to create a *superpod* (cf. Appendix A and §4.2). The second is to use the DCN network to scale out the capacity. These two fabrics complement each other in balancing compute and communication, scale and efficiency: the scale-up ICI within a superpod provides 50–100× more bandwidth than the DCN at a fraction of the cost; the scale-out DCN enables training extremely large ML models.

2.2.1 *Scaling-up within a superpod*

Using a lightwave fabric within a superpod leads to significant cost and power savings while adding new capabilities and flexibility not possible with a traditional datacenter network or with a directly-connected static ML topology:

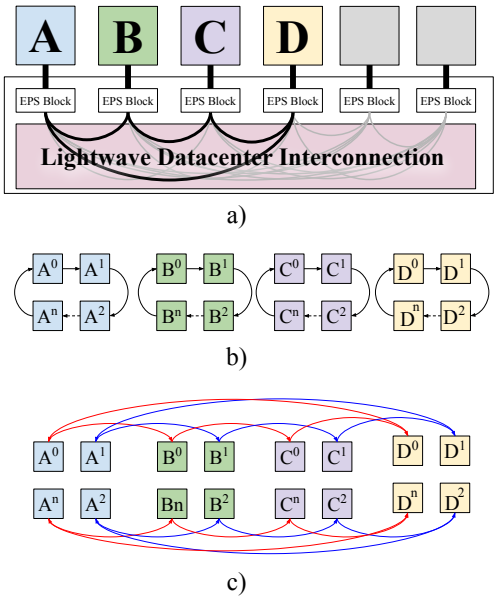

Figure 2: Scaling-out superpods using the DCN. a) A hybrid ICI-DCN network composed of four superpods connected by the DCN, b) A collective communication pattern within individual superpods using a ring topology and the ICI network, c) A collective communication between superpods using two rings (shown as red and blue) and the DCN network.

Reconfigurability: The cluster-level job scheduling system uses the lightwave fabric within a superpod to dynamically create workload-sized compute "slices" and program the inter-chip-interconnect network for the slice. This reconfigurability leads to the unique ability to control the configuration or shape of the slice to optimize performance (§4.2.1) and scheduling efficiency (§4.2.4).

Availability: The lightwave fabric provides significantly improved availability and overall throughput (§4.2.2).

Modular deployment: The use of lightwave fabric for a superpod enables increased deployment speed through modularity and the ability to incrementally deploy computing resources (§4.2.3).

2.2.2 *Scaling-out between superpods*

The superpod ICI network and the datacenter DCN network can be combined for ML training models that are too large to fit in a single superpod [12]. For this case, the lightwave fabrics in both the superpod ICI and the DCN are coordinated to create an interconnection in a hybrid ICI-DCN network as shown in Fig. 2 a). We optimize the workload end-to-end, starting from adapting the collective operations for the difference in ICI vs. DCN bandwidth per TPU, optimizing the topology within the pods, and cooptimizing job placement (TPU allocation) and reconfiguration of the DCN level topology to achieve high performance. Example collective communication patterns both within a pod using the ICI network and between pods using the DCN network are shown in Figs. 2 b) and c) respectively. While this optimization maximizes the communication within the ICI network, the transfers over the DCN network during c) are still on the critical path and delays can substantially affect the model throughput [20, 41]. The use of reconfigurable lightwave fabrics—both within a single superpod and within the DCN network—leads to substantial performance, capital and operational benefits for this kind of hybrid ICI-DCN network.

2.3 Requirements for the lightwave fabric

Ideally we want a lightwave fabric that can dynamically adapt its configuration to best suit the communication patterns of the workload. The fabric also needs to be performant, fault tolerant, and cost effective. These features lead to the following high-level requirements for lightwave fabrics:

Switch radix: A large switch radix is required to scale out to large deployments connecting hundreds of networking aggregation blocks for DCN and thousands of compute nodes for ML.

Reconfiguration flexibility: The OCS must be rapidly reconfigurable and non-blocking to dynamically control the topology. This includes the ability to keep certain connections undisturbed while making changes elsewhere. This requirement provides job isolation.

Transceiver performance: The transceivers for both DCN and ML applications must support a high-bandwidth bidirectional (bidi) link per fiber to reduce the overall system cost. For ML systems, the link must be low latency to support tight synchronization. For DCN networks, the transceiver must be backward compatible (i.e., each new generation must inter-operate with previous generations of transceivers).

Fabric Reliability: The fabric must be fault-tolerant at scale for high availability and robust performance in the presence of hardware and node failures.

Fabric cost effectiveness: The lightwave fabric must provide a total-cost-of-ownership (TCO) benefit.

Developing optical technologies to meet these requirements is a significant challenge—particularly at scale.

3 HARDWARE COMPONENTS

This section presents the in-house development of our optical switch and custom bidirectional (bidi) optical transceivers that enabled the wide variety of benefits of lightwave fabrics for DCN and ML applications.

3.1 Hardware Overview

The optical hardware developed for our lightwave fabrics has some common requirements for the DCN and ML use cases, as well as different requirements driven by the unique needs of our applications. Figure 3 shows the optical data path for the DCN and the ML use cases. Our current lightwave fabrics use bidirectional (bidi) transceivers in an OSFP form factor [42] and the Palomar OCSes within each OCS rack. This optical switch creates direct optical connections between the end points (North port (N) to South port (S) in figure). The OSFP module has two coarse WDM (CWDM4) transmitter/receiver pairs, each supporting four wavelength channels. The module is connected to two optical fibers with each fiber

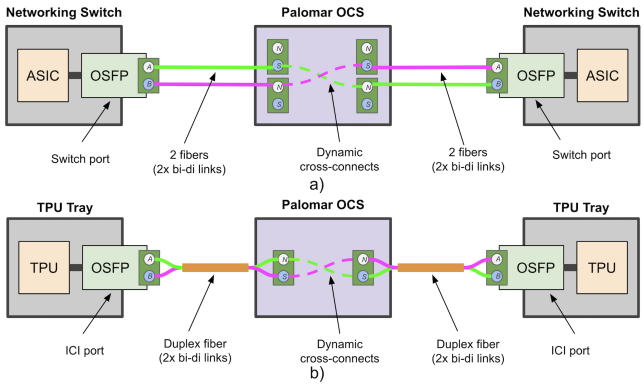

Figure 3: Components for the lightwave fabric for a) datacenter networking and b) ML superpod: bidi OSFP optical transceivers, duplex fiber with each fiber strand supporting a bidirectional link, and the Palomar OCS.

supporting a bidirectional (bidi) link. This means that each bidi transceiver supports twice the bandwidth compared to a standard duplex transceiver for which each fiber carries a single transmitted WDM signal or a single received WDM signal. For the DCN use case, two bidi links from each OSFP port may go to two different OCS ports for maximum fanout. For ML, two bidirectional links are bundled into a pair of fibers that connect to a duplex OCS port (N/S pair).

3.2 Development of the OCS

The heart of the lightwave fabric shown in Fig. 3 is the optical circuit switch. There are a wide variety of lightwave devices with different physical switching mechanisms [7, 46, 51, 55, 64] that can be used to construct an optical circuit switch. Appendix C compares cost, performance, and reliability/availability of several OCS technologies. Early use cases for OCS technology [40, 48, 64] recognized the utility of leveraging the CapEx development costs of OCS technology against the long-term value for OpEx in terms of network and bandwidth management to minimize the TCO. An additional benefit of OCS technology is that the required energy per switched bit can be orders of magnitude lower than EPS technologies because there is no per-packet processing.

3.2.1 Key Challenges

The choice of an appropriate OCS technology for a lightwave fabric is driven by the requirements outlined in § 2.3 and is constrained by the practical issues of cost, manufacturability, and reliability. Accordingly, this choice must consider:

Port count: For the DCN use case, the radix of the OCS and the network aggregation blocks determines the total number of aggregation blocks that can be supported, and thus maximum scale of the datacenter network. Similarly, for the ML use case, the radix of the OCS, size of an elemental compute building block, and the size of the routing table that can be supported determine the overall size of the TPU Superpod.

Switching time: Most optical switches have much slower switching times compared to electrical packet switches. They are therefore more suitable for topology engineering of persistent traffic [47], and/or creating custom topologies for predictable machine-learning workloads.

Optical performance: Optical link budget is a precious commodity for lightwave fabrics, greatly affecting cost and performance of the optical transceiver. The use of cost-effective, low-power optical transceivers with moderate link budget requires driving down the insertion loss through the OCS, ideally below 3dB. The requirement for low return loss (i.e., signal reflections back along the link) stems from the use of bidirectional links and is described in detail in § 3.3.

Low latency: The absence of per-packet processing within an OCS means only a small amount of deterministic latency is added on a per-hop basis, which is a key requirement for synchronous ML workloads. In comparison, other kinds of network fabrics that do not use direct connections can add hundreds of nanoseconds if not microseconds of delay per hop [30, 32].

Among the technologies listed in Table C.1 in Appendix C, MEMS OCS technology currently provides the best match for meeting the system-level challenges and the practical constraints of scale and economics for both the datacenter and ML use cases. The requirements of other use cases may dictate the use of other optical switching technologies and are discussed in § 6.

At the time when we first considered OCSes for datacenter applications, the demand for MEMS OCSes from traditional telecommunications applications was insufficient to support the volumes and availability required at our datacenter scale. The difficulties in maintaining reliability and quality of a vendor based OCS led to the decision to develop the **Palomar OCS**. Palomar has high availability, high reliability, is integrated into our networking control and monitoring infrastructure, and is used ubiquitously for the datacenter and ML use cases.

3.2.2 Optical Switch Design

Figure 4 shows the high-level optical design and operation principles of the Palomar OCS. The input/output (bidirectional) optical

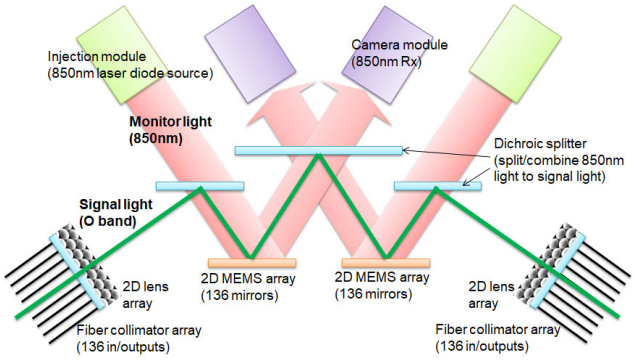

Figure 4: Illustration of the design and optical path of Palomar OCS optical core.

signals enter the optical core through two-dimensional (2D) fiber collimator arrays, which produce pencil-like beams that propagate through the switch. Each collimator array consists of a 136×136 fiber array and a 2D lens array. The optical core consists of two

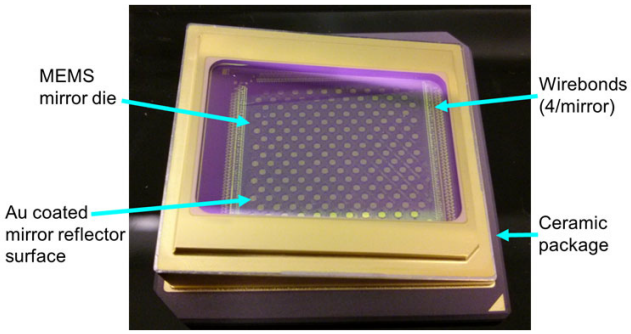

Figure 5: Photograph of a Palomar MEMS mirror package. Inside each ceramic package is a single large die with individually controllable micro-mirrors.

2D MEMS mirror arrays as shown in Fig. 5. To increase yield and redundancy, 176 micro-mirrors were fabricated on each MEMS die from which the best 136 mirrors were used for the switch with additional qualified connections used as manufacturing spares.

Each optical signal to be switched traverses through a port in each collimator array and the two MEMS mirror arrays, as indicated by the green line in Fig. 4. Mirrors on each array are actuated and tilted to switch the input signal to a corresponding input/output collimator fiber. The entire end-to-end optical path is broadband and reciprocal, for data-rate agnostic and bidirectional communication across the OCS. Figure 6 shows a photograph of the optical core and corresponding key components.

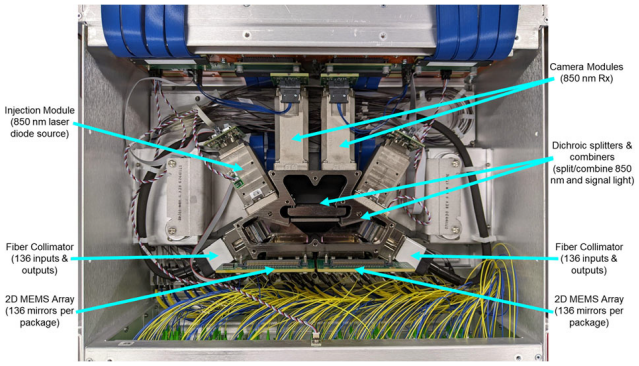

Figure 6: Photograph of Palomar OCS optical core with fiber collimators, camera modules, packaged MEMS arrays, injection modules, and dichroic splitters and combiners.

From a systems perspective, this design yields a non-blocking, 136×136 OCS with bijective, any-to-any input (North) to output (South) port connectivity. The non-blocking functionality provides important scheduling flexibility for both of our use cases. As an example, for the ML use case, slices for new model placements within a superpod can be dynamically scheduled without interfering with existing models running on a different slice. The benefits of this feature are discussed further in § 4.2.4.

A novel design choice that enabled us to realize a low-cost, manufacturable OCS was the use of two cameras, one per MEMS array (Fig. 4) for closed-loop alignment. The monitoring channel (red arrows, Fig. 4) for each camera is superposed with the signal path to be switched. Each MEMS array is thus illuminated with a 850nm monitoring beam. This monitor wavelength is different from the data-carrying signal wavelength which is around 1300 nm. The illuminated mirror arrays are then imaged onto the camera module using optical dichroic splitters that separate the monitor wavelength from the signal wavelength. The images from these cameras are used to provide feedback to optimize the position of each mirror for minimum loss. By implementing mirror controls based on image processing, the control scheme is significantly simplified compared to conventional approaches which can require individual per mirror monitoring and/or photodetector hardware.

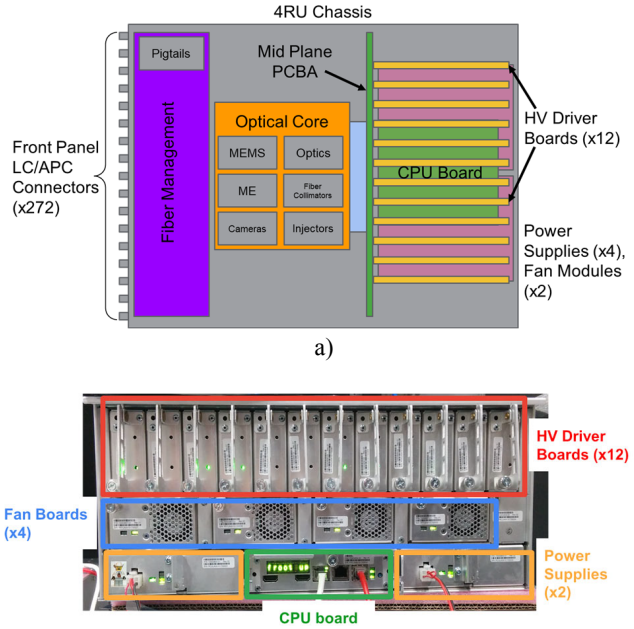

Figure 7: a) System-level diagram of the Palomar OCS showing the front half with fiber management and the optical core and the back chassis with the primary system boards. b) Rear view of the Palomar OCS back chassis showing the field replaceable units.

Several engineering principles guided the development of the Palomar OCS to achieve the necessary manufacturability, serviceability, and reliability needed for our systems. Fig. 7a) shows the system-level diagram of the OCS chassis. The CPU receives port connection commands from the control plane, which are then sent to an FPGA connected to a set of high voltage (HV) driver boards for mirror actuation. The front and back chassis architecture shown in Fig. 7 allows the optical core and primary printed circuit boards to be assembled and tested separately. This architecture enhances testability, yield, and also enables field serviceability of key components to increase availability. The power supplies and fan modules are redundant and can be hot swapped while maintaining functionality. Fig. 7b) shows a rear view of the OCS chassis. Although the mirror state cannot be maintained when driver boards are hot swapped, designing them to be field replaceable was critical as

the HV drivers for the mirrors was one of the largest reliability challenges for the switch.

We use the same software stack and base OS as our other datacenter networking devices (i.e., EPSes) for both control and *in-situ* evaluation of the state of the OCS. The commonality in management plane interfaces provides seamless integration with our existing infrastructure. We invested heavily in improving telemetry and anomaly reporting to account for the complexity of the hardware and the software interactions that manage it and the high reliability requirements. The ability to deeply integrate the control and monitoring software with the rest of our network infrastructure was essential given that the switches had a large "blast radius".

3.3 Development of Bidi Transceivers

For both our DCN and ML use cases, our studies indicated a large benefit in using bidi optical links. Achieving the combined goals of high bandwidth and bidi operation was a significant challenge—particularly at scale because these functionalities did not exist or were ahead of commercial standards-based transceivers and had to be developed by leveraging the economies of scale of datacom transceivers. The custom bidi WDM transceivers developed for our ML use case are significantly different compared to the standards-based point-to-point datacom transceivers [2] and are different from the transceivers developed for our DCN use case [59]. These differences result from supporting the high-bandwidth, low latency, synchronous nature of ML workloads. For both sets of transceivers, key components of the bidi transciever (laser, circulator, transimpedance amplifier (TIA), photodetector (PD), digital signal processing (DSP) ASIC, etc.) were designed or selected to emphasize module integration for low cost and manufacturability.

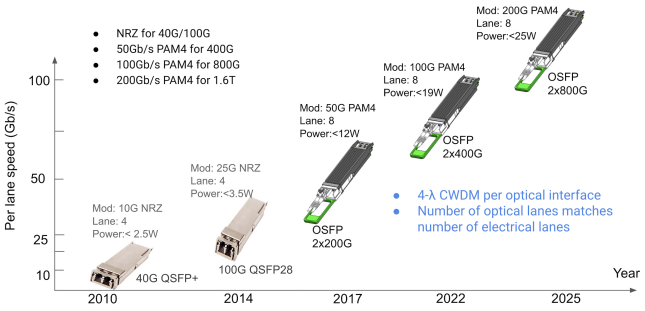

Figure 8: WDM interconnect for datacenter network.

Figure 8 shows the WDM interconnect products and roadmap for our datacenter network. From 40Gb/s QSFP+ to the latest 800Gb/s OSFP, the bandwidth of CWDM4 transceivers for DCN has grown 20× with continuous improvement in energy efficiency and linear density. Figure 9 shows the the latest generation bidi OSFP transceiver modules for use in ML superpods.

3.3.1 *Key challenges*

Consideration of the entire lightwave fabric for both ML and DCN use cases led to the following challenges for the bidi transceiver:

High bandwidth: As bandwidth requirements scaled in the datacenter, adoption of WDM was critical for the reach needed and

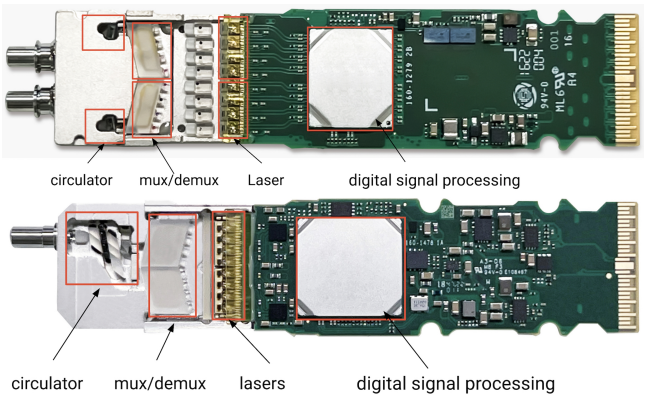

Figure 9: OSFP bidirectional modules. Top: bidi 2x 400Gb/s CWDM4 transceivers with 2 integrated circulators (cf. Appendix B). Bottom: bidi 800Gb/s CWDM8 transceiver with 8 multiplexed wavelength channels and a single integrated circulator.

efficient use of OCS technology. To enhance ML pod performance, a high bandwidth per fiber is critical to scale the lightwave fabric. Within the same spectral width (80nm) as a standard CWDM4 transceiver [2], we thus increased the number of wavelength lanes from 4 to 8 using a tighter wavelength spacing of 10nm as opposed to a standard 20nm spacing used for the CWDM4 bidi transceivers for datacenter networking and initial ML use cases.

Fiber impairments:

The transceivers for both the DCN use case (4×20nm) and the ML use case (8×10nm) operate over a 80 nm spectral range so that chromatic dispersion is an issue for data rates above 100 Gb/s for the link lengths used for our use cases. This impairment can be mitigated by managing frequency variations (chirp) in the laser and the modulator along with the use of nonlinear equalizers based on maximum likelihood sequence estimation (MLSE).

Larger optical link budget: To support the higher loss budget due to the OCS and circulators, low-loss optical components (thin-film-based wavelength mux/demux) and integrated circulators were used to minimize optical path loss. We use digital signal processing (DSP) to mitigate the impairment of optical multi-path-interference (MPI) and increase the overall link margin by using a novel concatenated forward-error-correction (FEC) technique. These two methods are discussed in § 3.3.2.

Bidirectional functionality: A key enabling technology which permitted OCS deployment to be cost-effective at scale was the development of optical circulator technology that permitted the bidirectional operation of the links. This feature effectively doubled the number of ports on the OCS.

When we started developing our transceivers, a few types of optical circulators were commonly employed in relatively limited quantities within telecommunications systems to increase the gain generated from a single optical amplifier, but there was a lack of a commercial high-volume ecosystem for this component. Similar to the other hardware technologies discussed in this paper, the baseline design of the circulators developed for telecommunications applications had to be re-engineered to optimize its performance

to meet our needs. These needs included operating in a different wavelength range, and reducing return loss and crosstalk between the ports. The crosstalk was particularly important as corresponding stray light is effectively equivalent to having a reflection in the link. Additional details on the operation and our use of optical circulators is provided in Appendix B.

Backward compatibility: A critical requirement to support smooth evolution for the DCN use case is backward compatibility of the optical transceiver technology. This requirement was achieved by a combination of careful design of the wavelength grid and analog front-ends (laser driver, TIA), developing programmable modules and DSP blocks that can run at multiple line rates along with the corresponding qualification testing for all supported rates. As example, the latest generation OSFP transceiver (cf. Fig. 8) running at 100G PAM4 per lane must also support 50G PAM4 and 25G NRZ operation. The mode of operation is software programmable, enabling inter-operation with the existing, legacy network fabric and the maximum rate when the system is upgraded.

3.3.2 *Digital Signal Processing ASIC*

State of the art digital signal processing (DSP) blocks [10] form the engine of our bidi WDM optical transceivers. The DSP not only provided a more robust, scalable solution by relaxing the requirements on the optical and analog electrical components, it also enabled new digital capabilities to mitigate optical impairments and increase the optical link budget. There are two custom DSP blocks implemented in our bidi transceivers: 1) Optical interference mitigation (OIM) [66], and 2) Concatenated forward error correction (FEC). The choice of these codes must balance more powerful (and higher-latency) codes that can improve the link margin with the requirement of low latency for ML workloads. The evaluation of these methods is given in §4.1.2.

4 EVALUATION

Meeting the stringent requirements on the optical hardware described in Section 3 yields a lightwave fabric with numerous benefits. This section evaluates the key features of the OCS and transceivers used to construct the lightwave fabrics and then evaluates the system-level benefits for our ML use case using a static non-reconfigurable fabric as the baseline.

4.1 Hardware Evaluation

The combined use of bidirectional links and OCSes requires careful evaluation of the optical components compared to a standard point-to-point link because of the potential interactions between the transceivers and the OCS.

4.1.1 *Optical Switch Evaluation*

Applying the design principles described in Section 3.2, tens of thousands of Palomar OCS with 136×136 duplex ports (each with eight spare ports) have been manufactured and deployed in our infrastructure. The maximum power consumption of the entire system is 108W, which is a fraction of the power of an EPS system with the same switch capacity because the OCS does no per-packet processing. Figure 10 shows some representative insertion loss and return loss data for the Palomar OCS that is deployed for both the DCN and ML use cases. Insertion losses are typically less than

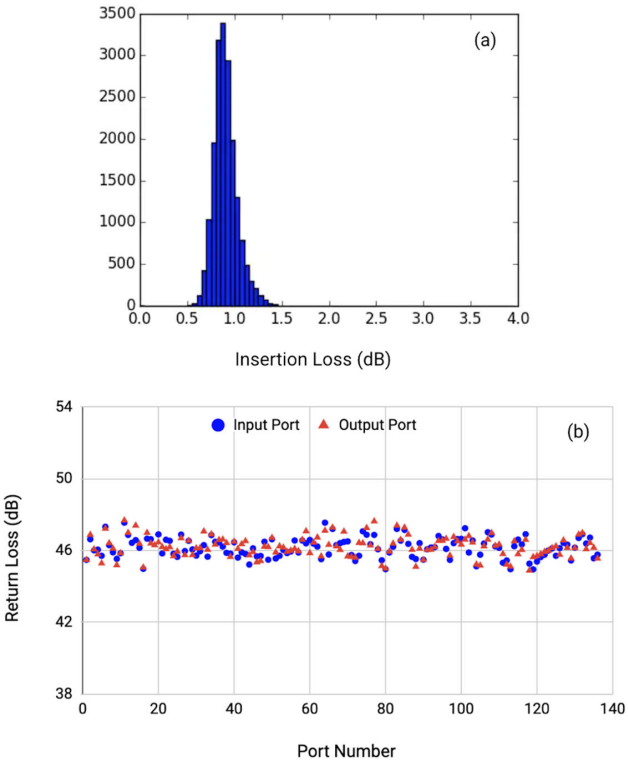

Figure 10: a) Representative Palomar OCS insertion loss histogram for 136×136 cross-connections. b) Return loss versus port number for 136 input/output ports.

2dB for all 136×136 permutations of connectivity. The tail in the distributions is nominally due to fiber splice and connector loss variation. Return loss caused by reflections is typically −46dB, with a nominal specification of less than −38dB. The major components of optical reflection come from the fiber collimators, at the interfaces between the fiber array and 2D lens array. This stringent return loss requirement stems from the use of bidirectional links along each optical path.

In terms of reliability and availability, the Palomar OCS passed all reliability tests (largely based on telecommunications/telcordia standards, i.e., damp heat, temperature cycle, shock and vibration, high/low temperature storage, etc.) with ample margins. For our production rack-mounted design, proper mechanical isolation of the optical core is critical due to vibration sensitivity of the MEMS mirrors and the need to maintain sub-micrometer level optical alignment. On-going reliability tests, manufacturing screens, and the ability to field replace failed sub-assemblies leads to the chassis typically achieving greater than 99.98% availability in the field today.

4.1.2 *Optical Transceiver Evaluation*

At the scale of millions of transceivers for all our applications, guaranteeing transceiver performance is a significant challenge as all corner cases in a high-dimensional parameter space including manufacturing, link quality, and component variations must be

effectively resolved.

Optical Interference Mitigation: Circulator-based bidirectional links have several unique physical impairments such as multi-path-interference (MPI) and in-band optical crosstalk. These impairments are caused by the reflections from optical devices within the path [15, 58, 60]. To mitigate MPI, a novel digital signal processing (DSP) based algorithm [66] has been developed. For this algorithm, the dominant carrier to carrier (interfering) beating noise, which exhibits a unique narrow-band spectral characteristic, is reconstructed in the digital domain and then removed from the received signal through a notch filter-based method. The center frequency of the notch filter is determined by monitoring the frequency offset between the source and the interfering carrier, also in the digital domain.

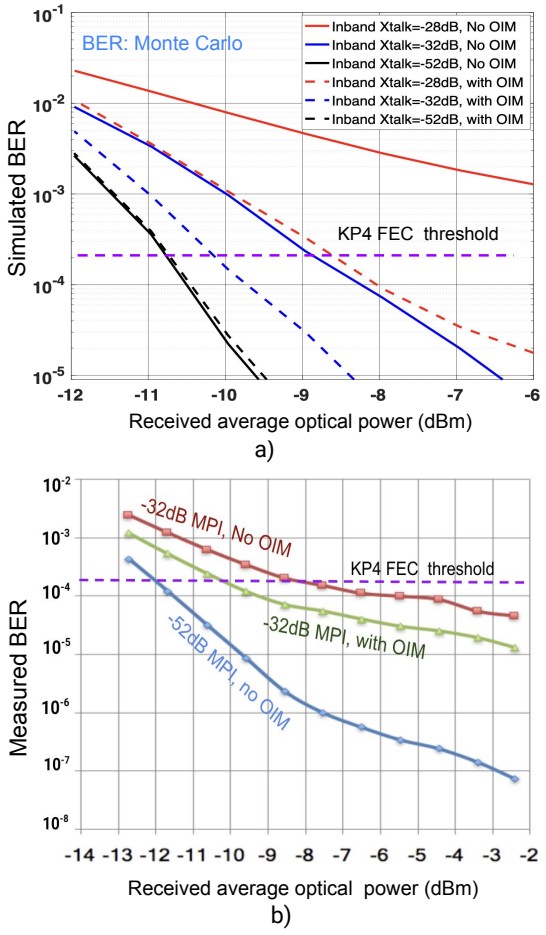

Figure 11: a) Simulated and b) measured receiver sensitivity in terms of bit error ratio for 50 Gb/s PAM4 for a single wavelength channel of a four channel (200 Gb/s) link.

This algorithm was first evaluated by modeling the optical link with and without the digital compensation. The link modeling incorporated all of the critical component parameters (bandwidth, noise, transmit power/receiver responsivity, insertion/return losses, extinction ratios, etc.). Figure 11a) shows the modeled effect on the bit error ratio (BER) for the link for several values of the MPI with and without optical interference mitigation (OIM) as shown by the dashed curves and solid curves, respectively. For example, for an MPI value of −32dB, and a bit error rate of 2×10^{-4}, which is the threshold for a standard KP4 error correcting code [22], the algorithm improves the receiver sensitivity by more than 1dB, which is significant for the total optical link budget.

In addition, the curves show the sensitivity of the BER to the MPI. This sensitivity clearly demonstrates the need for tight specification of components making up the lightwave fabric including the return loss of optical interfaces in the OCS and the circulators. Figure 11b) shows the corresponding measured data, which matches well with the modeling results and demonstrates the effectiveness of the OIM algorithm in mitigating multipath interference.

Concatenated Error Correction Coding: The DSP capability was also used to support a new ultra-low latency (<20ns for 200Gb/s) soft decision FEC (SFEC) code to increase the optical link margin. This proprietary code was used as an inner code and concatenated with a standard KP4 outer code [22]. A variant of this code has been adopted by the 200 Gb/s PAM4 IEEE Standards Group (IEEE802.3dj)[44].

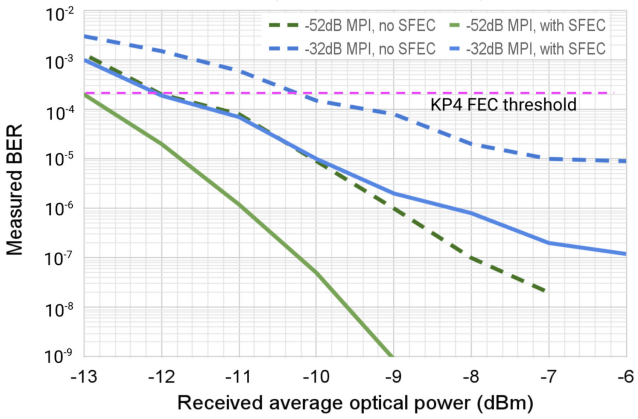

Figure 12: Optical receiver sensitivity improvement achieved by using only the concatenated soft-decision FEC (without OIM compensation) under two different MPI conditions.

Figure 12 shows the measured optical receiver sensitivity improvement achieved by concatenated SFEC under two different multipath interference (MPI) conditions. Consider the MPI value of −32dB with and without the new inner SFEC code as shown by the solid blue and dashed blue curves, respectively. At the KP4 outer code BER threshold of 2×10^{-4} (shown as the horizontal dashed magenta line), a large 1.6dB (45%) receiver sensitivity improvement can be achieved.

The most common metric to evaluate the system level performance of optical transceivers is the bit error rate (BER). Figure 13 shows a sampled distribution of the per lane BER data from a TPU V4 superpod in a span of days. The horizontal axis represents about 6144 (16 ports per cube face × 6 cube faces × 64 cubes) individual receiving ports, each paired with 64 possible other port partners (64

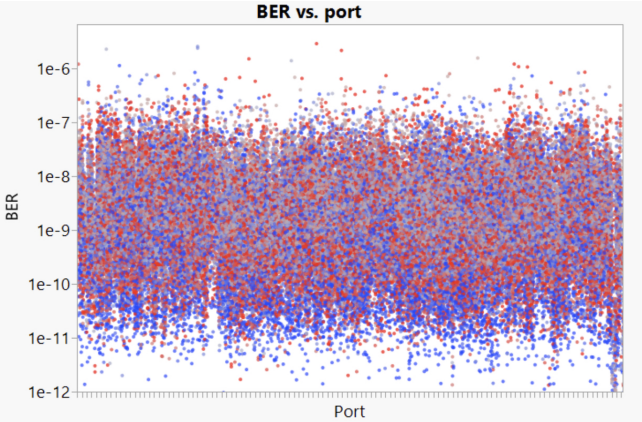

Figure 13: Bit-Error-Ratio (BER) (with OIM mitigation and SFEC) for production ML links versus port number.

× 6144 different 2× CWDM4 links). All of the values meet the KP4 error-correcting code specification of 2×10^{-4} with approximately two orders of magnitude of BER margin. This conservative BER margin validates the robustness of our design and leads to individual module failures that "are in the noise" compared to other failure modes that affect overall system availability.

4.2 ML Lightwave Fabric Evaluation

This section provides an evaluation of the system-level benefits of lightwave fabrics. We summarize our previous evaluation of the benefits of lightwave fabrics for DCN and provide a detailed evaluation of the benefits for the ML use case.

A detailed evaluation of lightwave fabrics in the datacenter use case was presented by Poutievski *et al.* [47]. That evaluation showed that a spine-free DCN delivers 30% reducion in CapEx and 40% reduction in OpEx compared to an spine-full Clos fabric that uses electrical packet switches. The CapEx and OpEx savings come from elimination of the spine layer and associated optical transceivers used in the switches. The use of a reconfigurable direct-connect lightwave fabric along with topology and traffic engineering also provides a 10% improvement in flow completion time and 30% increase in TCP throughput compared to a uniform mesh network.

The success of the DCN use case was the starting point for the development of the lightwave fabric for the TPU V4 superpod. For modularity, the TPU V4 superpod splits the intrapod interconnection fabric into an electrical part and an optical part as shown in Fig. 14. Electrical interconnects connect 64 TPU V4 chips to form an elemental cube with dimension $4 \times 4 \times 4 = 64$. The cube fits within a single rack and 64 cubes are then optically connected to the OCSes of the lightwave fabric to form a $64^2 = 4096$ TPU superpod (see Appendix A for further details about the architecture of TPU superpod).

The cluster-level job scheduling system uses the lightwave fabric to dynamically connect one or more elemental cubes to create compute "slices" and program the inter-chip-interconnect network for the slice. Most slices have a 3D torus topology with wrap-around links. The number of cubes connected per dimension of the torus can be different, e.g., when the entire pod of 4096 nodes is used,

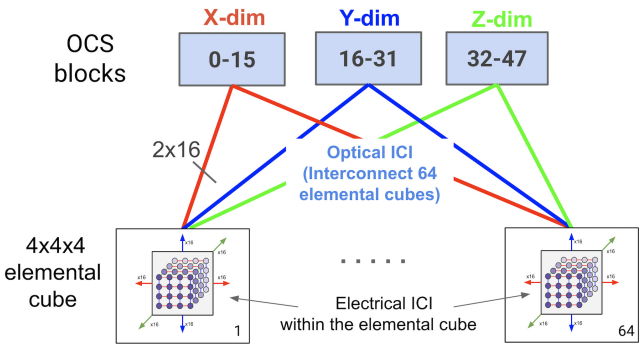

Figure 14: TPU V4 superpod. Each $4 \times 4 \times 4$ cube is statically connected using electrical interconnections and fits within a single rack. Optical inter-cube connections (colored lines) are flexibly reconfigured with the OCS.

	DCN	Lightwave Fabric	Static
Relative Cost	1.24X	1.06X	1X
Relative Power	1.10X	1.01X	1X

Table 1: Cost and power comparison for three different fabrics used to connect 4096 TPU V4 chips, normalized to a static topology.

slice topologies ranging from $4 \times 4 \times 256$ to $16 \times 16 \times 16$ can be configured with the minimum increment of four set by the size of the elemental $4 \times 4 \times 4$ cube. Multiple different-sized workloads with differing communication patterns can be configured within a single pod. The lightwave fabric enables each workload to run on a different slice that is physically isolated from the other jobs.

Table 1 shows the relative cost and power of a reconfigurable lightwave fabric compared to other networking fabrics that could be used to support a TPU V4 Superpod. Our baseline for this comparison is a static topology that uses short-range, low-cost optical interconnect to directly connect the 64 elemental cubes. Our internal studies indicate that the cost of a lightwave fabric is 6% more expensive and uses 1% more power compared to the baseline static fabric. For comparison, the same analysis shows that a superpod with our EPS-based DCN fabric is 24% more expensive and uses 10% more power.

Using a symmetric static fabric as our baseline, we evaluate four key benefits of a reconfigurable lightwave fabric for ML: 1) Execution speedup, 2) Fabric availability, 3) Increased deployment speed and flexibility and 4) Efficiency.

4.2.1 *Speedup benefits of reconfigurability*

The reconfigurability of the lightwave fabric enables the creation of slices and provides significant performance optimization for large-scale ML model training. In normal operation, the routing is deterministic and set by the slice configuration. The transport protocol is proprietary with the collective communication library built into the XLA execution engine [49]. For a full-scale superpod containing 4096 TPU V4 chips, the dynamic reconfigurablity of the lightwave fabric enables changing the shape of the slice to any configuration between a symmetric slice of $16 \times 16 \times 16$ to a highly-asymmetric slice of $4 \times 4 \times 256$. This unique capability is especially important for emerging transformer-based large language models

(LLMs) [45, 57]. Performance optimization for these LLMs is critical, because these LLMs are very expensive to train, using hundreds to tens-of-thousands of computing elements including TPUs or GPUs for months [12, 45]. Table 2 shows that the reconfigurability of the

Model	Model Size (# of Params)	Optimal Configuration	Relative speedup w.r.t. baseline
LLM0	35 Billion	$8 \times 16 \times 32$	$1.54X$
LLM1	70 Billion	$4 \times 4 \times 256$	$3.32X$
LLM2	150 Billion	$16 \times 16 \times 16$	$1X$

Table 2: Optimal slice configuration and relative speedup (compared to a static $16 \times 16 \times 16$ baseline) on training throughput (samples processed per second) for several LLMs, the ML model category usually with largest model size.

lightwave fabric unlocks more than a 3.3× performance uplift for production-scale LLMs, using an optimally configured 4096 node TPU V4 superpod compared to the static baseline $16 \times 16 \times 16$ configuration. The symmetric $16 \times 16 \times 16$ static configuration is chosen as the baseline because it has the highest bisection bandwidth among all possible static configurations of a 3D torus. The optimal lightwave fabric configuration, together with the model-specific parallelism configuration including model partitioning and pipelining, is determined automatically by a reinforcement-learning-based and hyperscale-hardware-optimized neural architecture search (NAS) system without human intervention [33].

Another important observation from Table 2 is that there is no "one-size-fits-all" optimal slice configuration for ML models. As can be seen in Table 2, the symmetric $16 \times 16 \times 16$ configuration of the lightwave fabric is optimal for some LLMs, while asymmetric configurations are optimal for others. When training on large-scale ML systems, LLMs rely on state-of-the-art model parallelism (including model pipelining) and data parallelism [52] to achieve good training speed. The amount of inherent model and data parallelism for an LLM determines the optimal slice configuration. In general, model size determines the amount of inherent model parallelism, while global batch size determines the amount of inherent data parallelism.

Concretely, LLM2 has large model size and large global batch size, which provides sufficient model and data parallelisms. Therefore, LLM2 prefers the $16 \times 16 \times 16$ cube slice configuration to leverage the maximum bisection bandwidth. For LLM0 and LLM1, they have much larger global batch size than their model size and thus much higher amount of inherent data parallelism than model parallelism. Therefore, both LLM0 and LLM1 prefer asymmetric configurations to match the inherent imbalanced parallelism on the model and the data dimensions, with LLM1 preferring a more asymmetric configuration because of its inherent parallelism being more skewed to data parallelism. When possible, our automated model parallelization optimizer assigns the 1st dimension of the TPU slice to the model parallelism and the 2nd and 3rd dimensions to the data parallelism. Thus, our optimizer finds the optimal slices for LLM0 and LLM1 as $8 \times 16 \times 32$ and $4 \times 4 \times 256$ with 1.54X and 3.32X speedup compared to the symmetric baseline of $16 \times 16 \times 16$, respectively. Our initial work in co-optimizing the model and the slice configuration demonstrates the unique and profound capability of lightwave fabrics in accelerating giant ML models at scale. More in-depth exploration in this space is the topic of future work.

The diverse dependence of ML model performance on the interconnect configuration requires the size and topology of the slice to be late binding after hardware is deployed. The lightwave fabric thus not only achieves substantial performance gains for ML models today but also offers a unique and powerful form of reconfigurability for future ML models at scale.

4.2.2 *Fabric Availability*

The ability to reliably compose slices is critically dependent on the availability of the lightwave fabric. As shown in Fig. 14, each building block used in the lightwave fabric is arranged as a $4 \times 4 \times 4$ cube with six faces with each face having $4^2 = 16$ connections. Each connection has 8 optical lanes, and Palomar OCS has a port count of 128×128. We would need 96 OCSes using standard CWDM4 duplex modules, but only 48 OCSes are required with our custom CWDM4 bidi module. While the current generation of superpod uses 48 OCSes with CWDM4 bidi modules, only 24 OCSes would be needed if we were to employ our custom CWDM8 bidi module. In all cases, a single failure in the set of OCSes that provide full connectivity between the elemental cubes will degrade the performance of any slice composed of more than one elemental cube.

The impact of OCS availability on the superpod lightwave fabric availability using three different kinds of optical transceivers is shown in Fig. 15a). The reduction in the required number of OCSes relaxes the availability requirement of a single OCS and improves the fabric availability from 90% using standard CWDM4 duplex to 95% using CWDM4 bidi, and 98% using CWDM8 bidi (assuming a single OCS availablity of 99.9%). Correspondingly, the overall effective throughput (i.e., goodput) for larger slices significantly improves compared to a static configuration over a wide range of host/server availability conditions.

To quantify this improvement, we use a single OCS availability of 99.9% and hold the overall system availability constant at 97%. For a single elemental cube slice (64 TPUs), to achieve this target availability requires holding back some elemental cubes within the pod so that there are enough working cubes to achieve the target system availability. The number of elemental cubes that are held back is directly proportional to the failure rate of an individual server. Larger failure rates decrease the goodput because more cubes need to be held back to achieve the target system availability. This can seen in Fig. 15b) for a single elemental cube slice (64 TPUs). As the server availability increases from 99% to 99.9%, the goodput increases because fewer elemental cubes need to be held back. For a slice that is a single cube, no reconfiguration between cubes is used and thus the goodput is the same for both the static and the reconfigurable topologies.

For slice sizes larger than a single cube (64 TPUs), the total number of cubes required for a reconfigurable lightwave fabric to achieve the target system availability is much smaller than the number of cubes required for a static fabric because the reconfigurable lightwave fabric can swap out a bad elemental cube whereas a static configuration cannot. For slices of the same size that are larger than 64 TPUs, this leads to the goodput for the static configuration (shown as dashed lines) rapidly degrading compared to

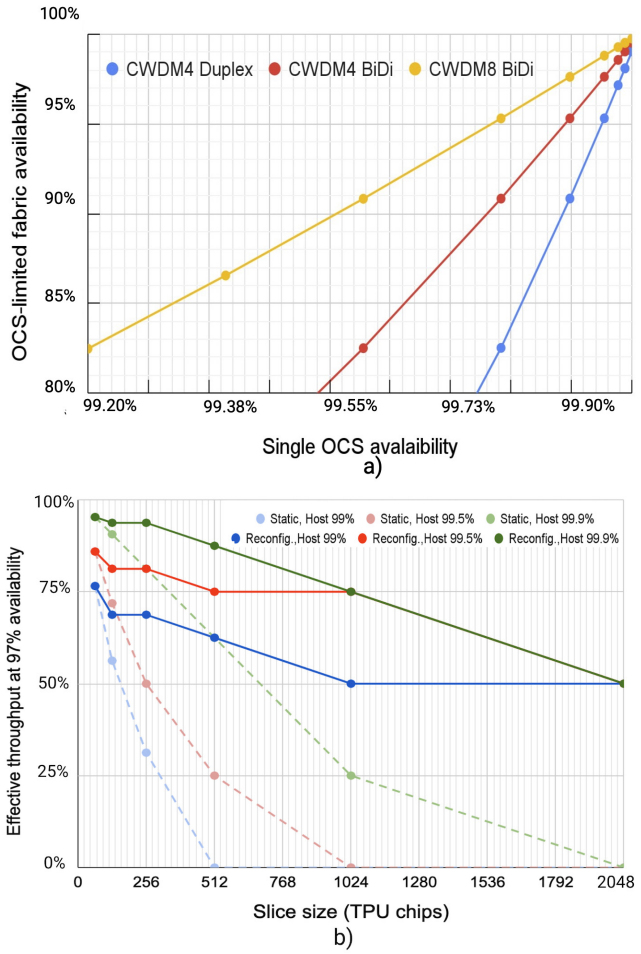

Figure 15: a) Impact of fabric availability on OCS availability for different transceiver technologies. b) Impact of server availability on the effective throughput of a superpod for a fixed overall system availability of 97% for different slice sizes.

the goodput for the corresponding reconfigurable lightwave fabric (solid lines). As an example, for a server availability of 99.9% (green curves), the static configuration can only support a 1024 TPU slice size with 25% goodput, whereas the reconfigurable superpod can support 1024 slice size with 75% goodput.

Given the simplifying constraint that the slices are the same size, as the slice size increases, the overall goodput is eventually dominated by the availability of the slice size and not by the individual server/host availability. At a slice size of 1024, this leads to the convergence of the goodput for a server availability of 99.9% (green curve) with the goodput for a server availability of 99.5% (red curve). Accounting for the hold back, three 1024 slices can be composed for either server availability leading to a goodput of 75% for both server/host availabilities. In contrast, only two 1024 slices with a goodput of 50% can be composed for the lower server availability of 99% (blue curve) because the hold back needed to achieve a 97% system availability for this case exceeds 1024 TPUs. At a slice size of 2048, which is half the pod size of 4096, only one slice can be composed—leading to a goodput of 50%—regardless of the server/host availability because some hold back is needed to achieve the target system availability for any server/host availability less than 100%. Accordingly, only one 2048 slice can be composed.

In practice, a distribution of slice sizes running different size models is used. The non-blocking nature of the OCS permits the scheduling of new slices of different sizes without interfering with slices that are already scheduled and running (c.f. § 4.2.4).

4.2.3 *Deployment Speed and Modularity*

The use of lightwave fabric for TPU superpods enables increased deployment speed through modularity and the ability to incrementally deploy racks containing an elemental cube. Because the inter-chip interconnect (ICI) for the 64 TPU chips is electrical and contained within a single rack, the connectivity and performance of each cube is verified when the chips and intrarack electrical interconnect is installed. The rack-level blocks can then be incrementally connected and verified at the pod level using the lightwave fabric to form larger slices up to the maximum of 64 blocks for a fully built-out pod. This modular, incremental deployment strategy significantly reduces the time to production and thus improves the cost effectiveness of lightwave-connected superpods compared to previous generations of superpods that did not use a lightwave fabric. For comparison, a TPU V3 superpod could not be verified until all 1024 chips and connecting cables were installed and tested.

The use of bidirectional transceivers also reduces the deployment time because only 48 OCSes (instead of 96 for standard CWDM4 transceivers) are needed. This saves 50% in the cost of the OCSes and fiber. Similar to the DCN use case, the associated CapEx costs can be amortized over the lifetime of the building (multi decade versus multi year).

4.2.4 *Efficiency benefits of reconfigurability*

The combination of an elemental cube with only 64 TPU chips and the reconfigurability enabled by the lightwave fabric leads to simplified slice scheduling, which increases overall efficiency. In the previous-generation TPU V3 superpod [26], scheduling a 256-node slice required finding 256 contiguous nodes that were idle and functional. For the TPU V4 superpod, the smaller 64-node block size in combination with the use of a non-blocking lightwave fabric means that there are many more potential solutions to configure the lightwave fabric to a connect a set of four idle, not-necessarily-contiguous $4 \times 4 \times 4$ elemental cubes to form a directly connected 256-node slice. In addition, the scheduler is able to defragment the pods more effectively. A detailed evaluation of this benefit is difficult to perform because it requires a large-scale controlled experiment to clearly isolate the benefits of reconfigurability. Nevertheless, in practice, we are able to run the TPU V4 fleet at a higher ($>$ 98%) utilization than earlier-generation superpods despite the need to support 4× larger slices. Details of the scheduling algorithm that achieves this utilization are the subject of a future paper.

5 RELATED WORK

The research community has proposed numerous lightwave datacenter network designs [8, 18, 31, 40, 50, 54, 56, 61, 62] and implemented technology demonstrators that incorporate OCSes into network architectures [16, 19, 50, 54, 61, 62]. The concept of topology engineering is prevalent through many of these works. The

work presented in this paper is distinguished from this related work by our development of the hardware and software that permits the production deployment of lightwave fabrics at scale.

Earlier system-level work and small-scale demonstrations of lightwave technology for HPC applications predate our work on TPU superpods [5, 9, 29, 50]. More recent studies discuss the potential benefits of topology management for HPC [38, 39] and machine-learning applications [31, 63]. Compared to this work, our use of lightwave fabric provides greater scalability, higher overall system availability, and greater performance by matching the configuration of the slices to the ML workloads.

Related work in transceiver design is extensive with many standards and multi-source agreements (MSA) [2] in place to reduce cost and allow interoperability between vendors. Our need to support bidirectional links and the loss through the OCS required us to divert from this path and develop custom transceiver designs for the DCN fabric and TPU superpod use cases. Related work on using circulators for increasing OCS and other optical switch radix has been proposed [54]. Our work advances this understanding by developing cost-effective, practical circulators [15, 58, 60].

6 FUTURE WORK

There are several lessons from our experience in developing lightwave fabrics that provide context for future work. Perhaps the most important is that at scale, "low-hanging fruit" can provide substantial benefits. The ability of lightwave fabrics to incrementally scale the system in a data rate agnostic manner and be considered as part of the building infrastructure provide sustained long-term benefits. The other key lesson is lightwave fabrics are by no means "one size fits all". The requirements for the lightwave fabric for the ML use case drove specific design choices that are distinct from the datacenter network use case. As a specific example, we did not know how the transceiver technology would evolve when we first started using lightwave fabrics at 40 Gb/s. Nevertheless, we have maintained interoperability across an order of magnitude difference in data rates (400 Gb/s vs. 40 Gb/s).

An important practical lesson is that at larger scales "everything breaks". Because it is increasingly difficult to test all system-level corner cases, this motivates the use of reconfigurable topologies enabled by lightwave fabrics that can be adapted to deal with unforeseen circumstances or workloads.

Looking forward to future lightwave fabrics, free-space MEMS OCS technology can be improved along all major performance axes (scale, switching time, loss), as evidenced by existing literature and our current internal development efforts to manufacture a larger 300×300 MEMS-based OCS with improved reliability and enhanced features for link-quality monitoring. Other switching technologies such as piezo-electric actuators [46] and silicon photonic MEMS [51] have fundamental advantages in faster switching time and lower drive voltages that could be more suitable for some applications.

One class of future use cases is based on long-lived or deterministic traffic patterns that can be supported by the existing and future OCS technologies. Each potential use case requires a specific codesign for the OCS, transceiver technology and/or switch/TPU chips. For datacenter networking, another potential use case is campus-wide networks. For ML, a different use case is supporting higher-dimensional topologies such as a 4D or 6D torus that has a larger bisection bandwidth, lower latency and greater scalability compared to a 3D torus.

A different class of future use cases are based on lightwave fabrics that can support faster reconfiguration times. For ML, changing the configuration of the slice during a training session to match communication patterns of different computing phases has the potential to improve performance [63]. Potential use cases for fast lightwave fabrics must balance the benefits with the challenge of developing transceivers with fast initialization times and sufficient link margin along with a control plane that can operate on the requisite time scale. Examples include technologies that focus on switching in the optical domain on nanosecond [4, 40] and microsecond-order timescales [37, 51].

The development of these new technologies can enable new use cases for lightwave fabrics that build upon our existing work (cf. §2.2.2) to create large-scale hierarchical hybrid electrical/optical networks for future large-scale workloads such as ML model training. This kind of hierarchical hybrid network is a promising path to support the demanding requirements of future large-scale workloads [12]. These networks will use electrical fabrics in concert with a co-designed combination of a lightwave fabric developed for slice management within a superpod and a different lightwave fabric developed for topology engineering within the data center network that connects the superpods.

7 CONCLUSION

This paper presented our experience in developing and deploying reconfigurable lightwave fabrics at scale. We note that while this paper focused on our initial development and deployment of free-space MEMS-based lightwave fabrics, our system-level architecture abstracts the underlying physical mechanisms of the OCS technology. As more lightwave technologies for both OCSes and optical transceivers mature and other use cases are developed, we can readily insert these emerging lightwave technologies into our existing system-level environments or develop new use cases based on nascent technologies that are appropriate for our needs.

Our experience shows that using practical constraints, performant, reliable and cost-effective production lightwave fabrics can be built and deployed at scale. We firmly believe that the networks described in this paper are just the first instance of a whole new class of reconfigurable network fabrics that are enabled by the rigorous co-design of optical circuit switching and transceiver technology.

ACKNOWLEDGEMENTS

We would like to acknowledge the Platforms Infrastructure Engineering (PIE), Palomar, Apollo, TPU Superpod, and ML system co-design & optimization (ML Scoop) teams, NetInfra, Global Capacity Delivery (GCD), Datacenter Operations, Site Reliability Engineering (SRE), Program Management Organization (PMO), and industry/vendor partners for making Lightwave Fabrics a reality at scale. We thank Safeen Huda for helping evaluate the speedup benefits of reconfigurability. We would also like to thank Parthasarathy Ranganathan, Shuang Yin and Alex Snoeren for helping review the manuscript.

REFERENCES

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2016. TensorFlow: A System for Large-Scale Machine Learning. In *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation (OSDI'16)*. USENIX Association, USA, 265–283.

[2] CW-WDM MSA (Continuous-Wave Wavelength Division Multiplexing Multi-Source Agreement). 2021. www.cwdm4-msa.org, Last accessed on 2023-1-30. (2021).

[3] Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. 2008. A Scalable, Commodity Data Center Network Architecture. In *Proceedings of the ACM SIGCOMM 2008 (SIGCOMM '08)*. Association for Computing Machinery, New York, NY, USA, 63–74. https://doi.org/10.1145/1402958.1402967

[4] Hitesh Ballani, Paolo Costa, Raphael Behrendt, Daniel Cletheroe, Istvan Haller, Krzysztof Jozwik, Fotini Karinou, Sophie Lange, Kai Shi, Benn Thomsen, and Hugh Williams. 2020. Sirius: A Flat Datacenter Network with Nanosecond Optical Switching. In *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM '20)*. Association for Computing Machinery, New York, NY, USA, 782–797. https://doi.org/10.1145/3387514.3406221

[5] K.J. Barker, A. Benner, R. Hoare, A. Hoisie, A.K. Jones, D.K. Kerbyson, D. Li, R. Melhem, R. Rajamony, E. Schenfeld, S. Shao, C. Stunkel, and P. Walker. 2005. On the Feasibility of Optical Circuit Switching for High Performance Computing Systems. In *SC '05: Proceedings of the 2005 ACM/IEEE Conference on Supercomputing*. Association for Computing Machinery, 16–16. https://doi.org/10.1109/SC.2005.48

[6] Brad Calder, Ju Wang, Aaron Ogus, Niranjan Nilakantan, Arild Skjolsvold, Sam McKelvie, Yikang Xu, Shashwat Srivastav, Jiesheng Wu, Huseyin Simitci, Jaidev Haridas, Chakravarthy Uddaraju, Hemal Khatri, Andrew Edwards, Vaman Bedekar, Shane Mainali, Rafay Abbasi, Arpit Agarwal, Mian Fahim ul Haq, Muhammad Ikram ul Haq, Deepali Bhardwaj, Sowmya Dayanand, Anitha Adusumilli, Marvin McNett, Sriram Sankaran, Kavitha Manivannan, and Leonidas Rigas. 2011. Windows Azure Storage: A Highly Available Cloud Storage Service with Strong Consistency. In *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles (SOSP '11)*. Association for Computing Machinery, New York, NY, USA, 143–157. https://doi.org/10.1145/2043556.2043571

[7] Calient Technologies. 2023. http://calient.net, Last accessed on 2023-1-30. (2023).

[8] Peirui Cao, Shizhen Zhao, Min Yee The, Yunzhuo Liu, and Xinbing Wang. 2021. TROD: Evolving From Electrical Data Center to Optical Data Center. In *2021 IEEE 29th International Conference on Network Protocols (ICNP)*. IEEE, 1–11. https://doi.org/10.1109/ICNP52444.2021.9651977

[9] R.D. Chamberlain, M.A. Franklin, and Ch'ng Shi Baw. 2002. Gemini: an optical interconnection network for parallel processing. *IEEE Transactions on Parallel and Distributed Systems* 13, 10 (2002), 1038–1055. https://doi.org/10.1109/TPDS.2002.1041880

[10] Frank Chang, Sudeep Bhoja, Jamal Riani, Ishwar Hosagrahar, Jennifer Wu, Sameer Herlekar, Arun Tiruvur, Pulkit Khandelwal, and Karthik Gopalakrishnan. 2016. Link Performance Investigation of Industry First 100G PAM4 IC Chipset with Real-time DSP for Data Center Connectivity, In Optical Fiber Communication Conference. *Optical Fiber Communication Conference*, Th1G.2. https://doi.org/10.1364/OFC.2016.Th1G.2

[11] Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, and Robert E. Gruber. 2006. Bigtable: A Distributed Storage System for Structured Data. In *7th USENIX Symposium on Operating Systems Design and Implementation (OSDI 06)*. USENIX Association, Seattle, WA. https://www.usenix.org/conference/osdi-06/bigtable-distributed-storage-system-structured-data

[12] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. 2022. PaLM: Scaling Language Modeling with Pathways. (2022). https://doi.org/10.48550/ARXIV.2204.02311

[13] James C. Corbett, Jeffrey Dean, Michael Epstein, Andrew Fikes, Christopher Frost, J. J. Furman, Sanjay Ghemawat, Andrey Gubarev, Christopher Heiser, Peter Hochschild, Wilson Hsieh, Sebastian Kanthak, Eugene Kogan, Hongyi Li, Alexander Lloyd, Sergey Melnik, David Mwaura, David Nagle, Sean Quinlan, Rajesh Rao, Lindsay Rolig, Yasushi Saito, Michal Szymaniak, Christopher Taylor, Ruth Wang, and Dale Woodford. 2013. Spanner: Google's Globally Distributed Database. *ACM Trans. Comput. Syst.* 31, 3, Article 8 (aug 2013), 22 pages. https://doi.org/10.1145/2491245

[14] Jeffrey Dean and Sanjay Ghemawat. 2008. MapReduce: Simplified Data Processing on Large Clusters. *Commun. ACM* 51, 1 (jan 2008), 107–113. https://doi.org/10.1145/1327452.1327492

[15] A. Farhood, B Smith, and S. Anderson. 2012. Improved MPI upper bound analysis. http://www.ieee802.org/3/bm/public/nov12/. (2012). Last accessed Jan. 2023.

[16] Nathan Farrington, George Porter, Sivasankar Radhakrishnan, Hamid Hajabdolali Bazzaz, Vikram Subramanya, Yeshaiahu Fainman, George Papen, and Amin Vahdat. 2010. Helios: A Hybrid Electrical/Optical Switch Architecture for Modular Data Centers. In *Proceedings of the ACM SIGCOMM 2010 Conference (SIGCOMM '10)*. Association for Computing Machinery, New York, NY, USA, 339–350. https://doi.org/10.1145/1851182.1851223

[17] Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung. 2003. The Google File System. *SIGOPS Oper. Syst. Rev.* 37, 5 (Oct. 2003), 29–43. https://doi.org/10.1145/1165389.945450

[18] Madeleine Glick, David G. Andersen, Michael Kaminsky, and Lily Mummert. 2009. Dynamically Reconfigurable Optical Links for High-Bandwidth Data Center Networks, In Optical Fiber Communication Conference and National Fiber Optic Engineers Conference. *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference*, OTuA3. https://doi.org/10.1364/OFC.2009.OTuA3

[19] Albert Greenberg, James R. Hamilton, Navendu Jain, Srikanth Kandula, Changhoon Kim, Parantap Lahiri, David A. Maltz, Parveen Patel, and Sudipta Sengupta. 2009. VL2: A Scalable and Flexible Data Center Network. *SIGCOMM Comput. Commun. Rev.* 39, 4 (Aug. 2009), 51–62. https://doi.org/10.1145/1594977.1592576

[20] Yanping Huang, Youlong Cheng, Ankur Bapna, Orhan Firat, Mia Xu Chen, Dehao Chen, HyoukJoong Lee, Jiquan Ngiam, Quoc V. Le, Yonghui Wu, and Zhifeng Chen. 2019. *GPipe: Efficient Training of Giant Neural Networks Using Pipeline Parallelism*. Curran Associates Inc., Red Hook, NY, USA.

[21] R. Hui. 2019. *Introduction to Fiber-Optic Communications*. Elsevier Science.

[22] IEEE 802.3cd Working Group. 2018. https://ieee802.org/3/cd/public, Last accessed on 2023-2-2. (2018).

[23] Telescent Inc. 2023. www.telescent.com/products, Last accessed on 2023-6-30. (2023).

[24] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hölzle, Stephen Stuart, and Amin Vahdat. 2013. B4: Experience with a Globally-Deployed Software Defined Wan. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM (SIGCOMM '13)*. Association for Computing Machinery, New York, NY, USA, 3–14. https://doi.org/10.1145/2486001.2486019

[25] Norm Jouppi, George Kurian, Sheng Li, Peter Ma, Rahul Nagarajan, Lifeng Nai, Nishant Patil, Suvinay Subramanian, Andy Swing, Brian Towles, Clifford Young, Xiang Zhou, Zongwei Zhou, and David A Patterson. 2023. TPU v4: An Optically Reconfigurable Supercomputer for Machine Learning with Hardware Support for Embeddings. In *Proceedings of the 50th Annual International Symposium on Computer Architecture (ISCA '23)*. Association for Computing Machinery, New York, NY, USA, Article 82, 14 pages. https://doi.org/10.1145/3579371.3589350

[26] Norman P. Jouppi, Doe Hyun Yoon, Matthew Ashcraft, Mark Gottscho, Thomas B. Jablin, George Kurian, James Laudon, Sheng Li, Peter Ma, Xiaoyu Ma, Thomas Norrie, Nishant Patil, Sushma Prasad, Cliff Young, Zongwei Zhou, and David Patterson. 2021. Ten Lessons From Three Generations Shaped Google's TPUv4i : Industrial Product. In *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)*. 1–14. https://doi.org/10.1109/ISCA52012.2021.00010

[27] Norman P. Jouppi, Doe Hyun Yoon, George Kurian, Sheng Li, Nishant Patil, James Laudon, Cliff Young, and David Patterson. 2020. A Domain-Specific Supercomputer for Training Deep Neural Networks. *Commun. ACM* 63, 7 (June 2020), 67–78. https://doi.org/10.1145/3360307

[28] Norman P. Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, Rick Boyle, Pierre-luc Cantin, Clifford Chao, Chris Clark, Jeremy Coriell, Mike Daley, Matt Dau, Jeffrey Dean, Ben Gelb, Tara Vazir Ghaemmaghami, Rajendra Gottipati, William Gulland, Robert Hagmann, C. Richard Ho, Doug Hogberg, John Hu, Robert Hundt, Dan Hurt, Julian Ibarz, Aaron Jaffey, Alek Jaworski, Alexander Kaplan, Harshit Khaitan, Daniel Killebrew, Andy Koch, Naveen Kumar, Steve Lacy, James Laudon, James Law, Diemthu Le, Chris Leary, Zhuyuan Liu, Kyle Lucke, Alan Lundin, Gordon MacKean, Adriana Maggiore, Maire Mahony, Kieran Miller, Rahul Nagarajan, Ravi Narayanaswami, Ray Ni, Kathy Nix, Thomas Norrie, Mark Omernick, Narayana Penukonda, Andy Phelps, Jonathan Ross, Matt Ross, Amir Salek, Emad Samadiani, Chris Severn, Gregory Sizikov, Matthew Snelham, Jed Souter, Dan Steinberg, Andy Swing, Mercedes Tan, Gregory Thorson, Bo Tian, Horia Toma, Erick Tuttle, Vijay Vasudevan, Richard Walter, Walter Wang, Eric Wilcox, and Doe Hyun Yoon. 2017. In-Datacenter Performance Analysis of a Tensor Processing Unit. In *Proceedings of the 44th Annual International Symposium*

on Computer Architecture (ISCA '17). Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3079856.3080246

[29] Shoaib Kamil, Ali Pinar, Daniel Gunter, Michael Lijewski, Leonid Oliker, and John Shalf. 2007. Reconfigurable Hybrid Interconnection for Static and Dynamic Scientific Applications. In *Proceedings of the 4th International Conference on Computing Frontiers (CF '07)*. Association for Computing Machinery, New York, NY, USA, 183–194. https://doi.org/10.1145/1242531.1242559

[30] M. R. Siavash Katebzadeh, Paolo Costa, and Boris Grot. 2020. Evaluation of an InfiniBand Switch: Choose Latency or Bandwidth, but Not Both. In *2020 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*. 180–191. https://doi.org/10.1109/ISPASS48437.2020.00033

[31] Mehrdad Khani, Manya Ghobadi, Mohammad Alizadeh, Ziyi Zhu, Madeleine Glick, Keren Bergman, Amin Vahdat, Benjamin Klenk, and Eiman Ebrahimi. 2021. SiP-ML: High-Bandwidth Optical Network Interconnects for Machine Learning Training. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference (SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 657–675. https://doi.org/10.1145/3452296.3472900

[32] Ang Li, Shuaiwen Leon Song, Jieyang Chen, Jiajia Li, Xu Liu, Nathan R. Tallent, and Kevin J. Barker. 2020. Evaluating Modern GPU Interconnect: PCIe, NVLink, NV-SLI, NVSwitch and GPUDirect. *IEEE Transactions on Parallel and Distributed Systems* 31, 1 (2020), 94–110. https://doi.org/10.1109/TPDS.2019.2928

[33] Sheng Li, Garrett Andersen, Tao Chen, Liqun Cheng, Julian Grady, Da Huang, Quoc V. Le, Andrew Li, Xin Li, Yang Li, Chen Liang, Yifeng Lu, Yun Ni, Ruoming Pang, Mingxing Tan, Martin Wicke, Gang Wu, Shengqi Zhu, Parthasarathy Ranganathan, and Norman P. Jouppi. 2023. Hyperscale Hardware Optimized Neural Architecture Search. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 3 (ASPLOS 2023)*. Association for Computing Machinery, New York, NY, USA, 343–358. https://doi.org/10.1145/3582016.3582049

[34] Hong Liu. 2021. 200G per lane for 800G and beyond. In *Workshop in 2021 Optical Fiber Communications Conference and Exhibition (OFC)*.

[35] Hong Liu, Cedric F. Lam, and Chris Johnson. 2010. Scaling Optical Interconnects in Datacenter Networks Opportunities and Challenges for WDM. In *2010 18th IEEE Symposium on High Performance Interconnects*. 113–116. https://doi.org/10.1109/HOTI.2010.15

[36] Hong Liu, Ryohei Urata, and Amin Vahdat. 2012. Optical Interconnects for Scale-out Data Centers. In *Optical Interconnects for Future Datacenter Networks*. Springer, New York, Chapter 2, 17–31.

[37] William M. Mellette, Rob McGuinness, Arjun Roy, Alex Forencich, George Papen, Alex C. Snoeren, and George Porter. 2017. RotorNet: A Scalable, Low-Complexity, Optical Datacenter Network. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM '17)*. Association for Computing Machinery, New York, NY, USA, 267–280. https://doi.org/10.1145/3098822.3098838

[38] Cyriel Minkenberg, German Rodriguez, Bogdan Prisacari, Laurent Schares, Philip Heidelberger, Dong Chen, and Craig Stunkel. 2015. Large-scale system partitioning using OCS. In *2015 International Conference on Photonics in Switching (PS)*. 235–237. https://doi.org/10.1109/PS.2015.7329011

[39] Cyriel Minkenberg, German Rodriguez, Bogdan Prisacari, Laurent Schares, Philip Heidelberger, Dong Chen, and Craig Stunkel. 2016. Performance Benefits of Optical Circuit Switches for Large-Scale Dragonfly Networks, In Optical Fiber Communication Conference. *Optical Fiber Communication Conference*, W3J.3. https://doi.org/10.1364/OFC.2016.W3J.3

[40] B. Mukherjee, I. Tomkos, M. Tornatore, P. Winzer, and Y. Zhao. 2020. *Springer Handbook of Optical Networks*. Springer International Publishing. https://books.google.com/books?id=EisDEAAAQBAJ

[41] Deepak Narayanan, Aaron Harlap, Amar Phanishayee, Vivek Seshadri, Nikhil R. Devanur, Gregory R. Ganger, Phillip B. Gibbons, and Matei Zaharia. 2019. PipeDream: Generalized Pipeline Parallelism for DNN Training. In *Proceedings of the 27th ACM Symposium on Operating Systems Principles (SOSP '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3341301.3359646

[42] OSFP-XD MSA. 2017. https://osfpmsa.org/specification.html, Last accessed on 2023-2-3. (2017).

[43] G.C. Papen and R.E. Blahut. 2019. *Lightwave communications*. Cambridge University Press.

[44] Lenin Patra, Arash Farhood, Rajesh Radhamohan, Will Bliss, Sridhar Ramesh, and Dave Cassan. 2023. FEC baseline proposal for 200Gb/s per Lane IM-DD Optical PMDs. https://www.ieee802.org/3/dj/public. (2023). Last accessed June 2023.

[45] David A. Patterson, Joseph Gonzalez, Quoc V. Le, Chen Liang, Lluis-Miquel Munguia, Daniel Rothchild, David R. So, Maud Texier, and Jeff Dean. 2021. Carbon Emissions and Large Neural Network Training. abs/2104.10350 (2021). arXiv:2104.10350 https://arxiv.org/abs/2104.10350

[46] Huber+Suhner Polatis. 2023. polatis.com, Last accessed on 2023-6-30. (2023).

[47] Leon Poutievski, Omid Mashayekhi, Joon Ong, Arjun Singh, Mukarram Tariq, Rui Wang, Jianan Zhang, Virginia Beauregard, Patrick Conner, Steve Gribble, Rishi Kapoor, Stephen Kratzer, Nanfang Li, Hong Liu, Karthik Nagaraj, Jason Ornstein, Samir Sawhney, Ryohei Urata, Lorenzo Vicisano, Kevin Yasumura, Shidong Zhang, Junlan Zhou, and Amin Vahdat. 2022. Jupiter Evolving: Transforming Google's Datacenter Network via Optical Circuit Switches and Software-Defined Networking. In *Proceedings of the ACM SIGCOMM 2022 Conference (SIGCOMM '22)*. Association for Computing Machinery, New York, NY, USA, 66–85. https://doi.org/10.1145/3544216.3544265

[48] R. Ryf, J. Kim, J.P. Hickey, A. Gnauck, D. Carr, F. Pardo, C. Bolle, R. Frahm, N. Basavanhally, C. Yoh, D. Ramsey, R. Boie, R. George, J. Kraus, C. Lichtenwalner, R. Papazian, J. Gates, H.R. Shea, A. Gasparyan, V. Muratov, J.E. Griffith, J.A. Prybyla, S. Goyal, C.D. White, M.T. Lin, R. Ruel, C. Nijander, S. Arney, D.T. Neilson, D.J. Bishop, P. Kolodner, S. Pau, C. Nuzman, A. Weis, B. Kumar, D. Lieuwen, V. Aksyuk, D.S. Greywall, T.C. Lee, H.T. Soh, W.M. Mansfield, S. Jin, W.Y. Lai, H.A. Huggins, D.L. Barr, R.A. Cirelli, G.R. Bogart, K. Teffeau, R. Vella, H. Mavoori, A. Ramirez, N.A. Ciampa, F.P. Klemens, M.D. Morris, T. Boone, J.Q. Liu, J.M. Rosamilia, and C.R. Giles. 2001. 1296-port MEMS transparent optical crossconnect with 2.07 petabit/s switch capacity. In *OFC 2001. Optical Fiber Communication Conference and Exhibit. Technical Digest Postconference Edition (IEEE Cat. 01CH37171)*, Vol. 4. PD28–PD28.

[49] Amit Sabne. 2020. XLA : Compiling Machine Learning for Peak Performance. (2020).

[50] L. Schares, X. J. Zhang, R. Wagle, D. Rajan, P. Selo, S. P. Chang, J. Giles, K. Hildrum, D. Kuchta, J. Wolf, and E. Schenfeld. 2009. A reconfigurable interconnect fabric with optical circuit switch and software optimizer for stream computing systems, In Optical Fiber Communication Conference and National Fiber Optic Engineers Conference. *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference*, OTuA1. https://doi.org/10.1364/OFC.2009.OTuA1

[51] Tae Joon Seok, Jianheng Luo, Zhilei Huang, Kyungmok Kwon, Johannes Henriksson, John Jacobs, Lane Ochikubo, Richard S. Muller, and Ming C. Wu. 2019. Silicon photonic wavelength cross-connect with integrated MEMS switching. *APL Photonics* 4, 10 (2019), 100803. https://doi.org/10.1063/1.5120063

[52] Mohammad Shoeybi, Mostofa Patwary, Raul Puri, Patrick LeGresley, Jared Casper, and Bryan Catanzaro. 2020. Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism. https://arxiv.org/abs/1909.08053. (2020).

[53] Arjun Singh, Joon Ong, Amit Agarwal, Glen Anderson, Ashby Armistead, Roy Bannon, Seb Boving, Gaurav Desai, Bob Felderman, Paulie Germano, Anand Kanagala, Jeff Provost, Jason Simmons, Eiichi Tanda, Jim Wanderer, Urs Hölzle, Stephen Stuart, and Amin Vahdat. 2015. Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network *(SIGCOMM '15)*. Association for Computing Machinery, New York, NY, USA, 183–197. https://doi.org/10.1145/2785956.2787508

[54] Ankit Singla, Atul Singh, Kishore Ramachandran, Lei Xu, and Yueping Zhang. 2010. Proteus: A Topology Malleable Data Center Network. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks (Hotnets-IX)*. Association for Computing Machinery, New York, NY, USA, Article 8, 6 pages. https://doi.org/10.1145/1868447.1868455

[55] Shunichi Sohma, Toshio Watanabe, Tomohiro Shibata, and Hiroshi Takahashi. 2005. Compact and Low Power Consumption 16×16 Optical Matrix Switch with Silica-Based PLC Technology, In Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference. *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference*, OThV4. https://opg.optica.org/abstract.cfm?URI=OFC-2005-OThV4

[56] Min Yee Teh, Shizhen Zhao, Peirui Cao, and Keren Bergman. 2020. COUDER: Robust Topology Engineering for Optical Circuit Switched Data Center Networks. (2020). https://doi.org/10.48550/ARXIV.2010.00090

[57] Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, YaGuang Li, Hongrae Lee, Huaixiu Steven Zheng, Amin Ghafouri, Marcelo Menegali, Yanping Huang, Maxim Krikun, Dmitry Lepikhin, James Qin, Dehao Chen, Yuanzhong Xu, Zhifeng Chen, Adam Roberts, Maarten Bosma, Vincent Zhao, Yanqi Zhou, Chung-Ching Chang, Igor Krivokon, Will Rusch, Marc Pickett, Pranesh Srinivasan, Laichee Man, Kathleen Meier-Hellstern, Meredith Ringel Morris, Tulsee Doshi, Renelito Delos Santos, Toju Duke, Johnny Soraker, Ben Zevenbergen, Vinodkumar Prabhakaran, Mark Diaz, Ben Hutchinson, Kristen Olson, Alejandra Molina, Erin Hoffman-John, Josh Lee, Lora Aroyo, Ravi Rajakumar, Alena Butryna, Matthew Lamm, Viktoriya Kuzmina, Joe Fenton, Aaron Cohen, Rachel Bernstein, Ray Kurzweil, Blaise Aguera-Arcas, Claire Cui, Marian Croak, Ed Chi, and Quoc Le. 2022. LaMDA: Language Models for Dialog Applications. (2022). https://arxiv.org/abs/2201.08239

[58] Ryohei Urata and Hong Liu. 2016. Datacenter interconnect and networking: Present state to future challenges.. In *IEEE Optical Interconnects Conference*.

[59] Ryohei Urata, Hong Liu, Kevin Yasumura, Erji Mao, Jill Berger, Xiang Zhou, Cedric Lam, Roy Bannon, Darren Hutchinson, Daniel Nelson, Leon Poutievski, Arjun Singh, Joon Ong, and Amin Vahdat. 2022. Mission Apollo: Landing Optical Circuit Switching at Datacenter Scale. (2022). https://arxiv.org/abs/2208.10041

[60] Ryohei Urata, Hong Liu, Xiang Zhou, and Amin Vahdat. 2017. Datacenter Interconnect and Networking: from Evolution to Holistic Revolution, In Optical Fiber Communication Conference. *Optical Fiber Communication Conference*, W3G.1.

https://doi.org/10.1364/OFC.2017.W3G.1
[61] Amin Vahdat, Hong Liu, Xiaoxue Zhao, and Chris Johnson. 2011. The Emerging Optical Data Center, In Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2011. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2011*, OTuH2. https://doi.org/10.1364/OFC.2011.OTuH2
[62] Guohui Wang, David G. Andersen, Michael Kaminsky, Konstantina Papagiannaki, T.S. Eugene Ng, Michael Kozuch, and Michael Ryan. 2010. C-Through: Part-Time Optics in Data Centers. In *Proceedings of the ACM SIGCOMM 2010 Conference (SIGCOMM '10)*. Association for Computing Machinery, New York, NY, USA, 327–338. https://doi.org/10.1145/1851182.1851222
[63] Weiyang Wang, Moein Khazraee, Zhizhen Zhong, Manya Ghobadi, Zhihao Jia, Dheevatsa Mudigere, Ying Zhang, and Anthony Kewitsch. 2023. TopoOpt: Co-optimizing Network Topology and Parallelization Strategy for Distributed Training Jobs. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. USENIX Association, Boston, MA, 739–767. https://www.usenix.org/conference/nsdi23/presentation/wang-weiyang
[64] Ming C. Wu, Olav Solgaard, and Joseph E. Ford. 2006. Optical MEMS for Lightwave Communication. *Journal of Lightwave Technology* 24, 12 (Dec 2006), 4433–4454. https://opg.optica.org/jlt/abstract.cfm?URI=jlt-24-12-4433
[65] S. J. Ben Yoo. 2006. Optical Packet and Burst Switching Technologies for the Future Photonic Internet. *Journal of Lightwave Technology* 24, 12 (2006), 4468–4492. https://doi.org/10.1109/JLT.2006.886060
[66] X. Zhou, R. Urata, E. Mao, H. Liu, and C. L. Johnson. 2016. In-band optical interference mitigation methods for direct detection optical communication systems. (2016). https://patents.google.com/patent/US10084547B2/en US Patent 10084547B2.

Appendices

Appendices are supporting material that has not been peer-reviewed.

A TPU SUPERCOMPUTER

This appendix provides an overview of the architecture for the TPU V4 chip and is adapted from [25]. As background, we have developed several generations of TPU chips [28] and deployed different supercomputer architectures [27] based on these chips with the most recent version of the chip being TPU V4 [26]. When considering the appropriate architecture for the TPU V4 chip, we wanted to scale up the number of chips by 4× versus TPU V3 just as TPU V3 was 4× TPU V2. Given the distance between TPU V3 racks, some wrap-around links required for the existing 2D torus topology were so long that they had to be optical to meet the reach requirement. Optical links are more than ten times more expensive than electrical links. At 4× the scale, there would be even more optical links. Moreover, there were concerns about the bisection bandwidth of a large 2D torus and the availability of a larger scale system due to host failure rates.

Based on our 4× scaling goal, and the concerns about bisection bandwidth and availability of a large 2D torus, we choose to use a 3D torus to increase the bisection bandwidth and use a directly-connected lightwave fabric. This fabric dynamically connects a set of smaller elemental building blocks with static electrical interconnections within each block. Because a 3D cube provides the best bisection bandwidth, this suggested an elemental block size of either 4×4×4 (64 chips) or 8×8×8 (512 chips). Because an elemental cube of 512 chips would require static electrical interconnects spanning multiple racks, a 4×4×4 (64 chip) building block was chosen with 4 TPU V4 processors per CPU host. Each CPU host has a datacenter network (DCN) connection. The 16 CPUs and 64 TPU V4s are all housed within one rack.

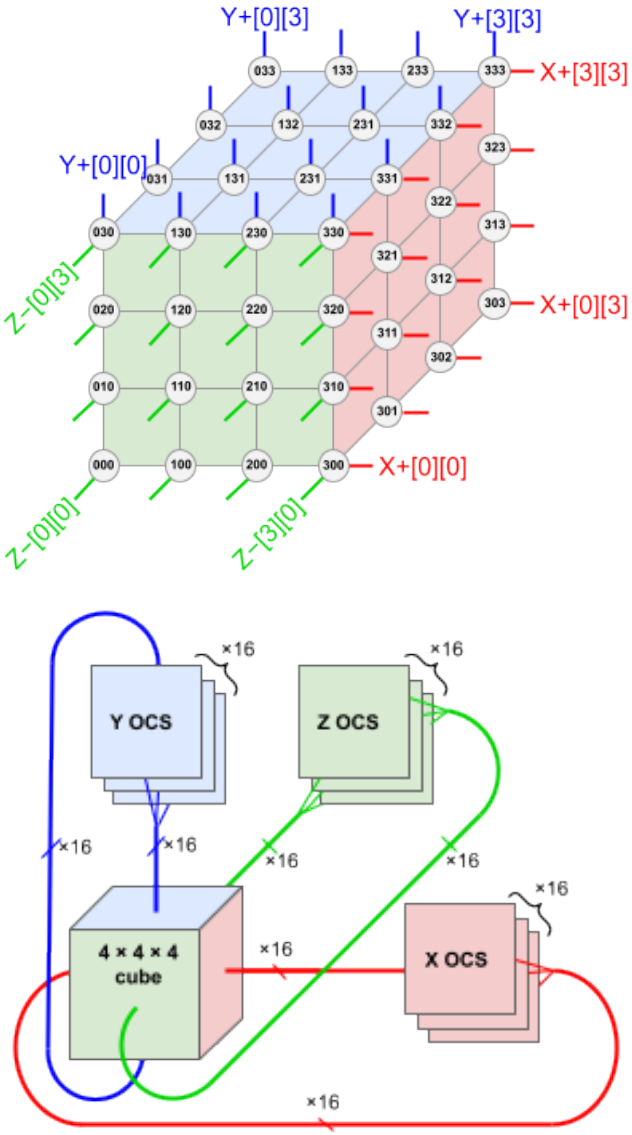

Figure A.1: Connectivity of a 4×4×4 cube (top) to 3 groups of OCSes (bottom). The "+" and "−" connections for each index and dimension are connected to the same OCS (from [25]).

Figure A.1 shows the links from the 6 "faces" of the basic 4×4×4 elemental cube used to configure the overall interconnect for the TPU superpod. There are 16 optical links per face, totaling 96 optical links per block that connect to OCSes. To provide the wraparound links to complete the 3D torus, the links on the opposing sides of a block are connected to the same OCS. Thus, each 4×4×4 block connects to $6 \times 16 \div 2 = 48$ OCSes. The Palomar OCS (described in § 3.2) is 136×136 (128 ports plus 8 spares for link testing and repairs), so 48 OCSes connect the 48 pairs of cables from 64 4×4×4 blocks, with each block composed of 64 chips. This configuration provides a total of $64^2 = 4096$ TPU V4 chips that can be connected using a

combination of an electrical interconnect within an elemental cube and a dynamic lightwave fabric between the blocks.

B OPTICAL CIRCULATOR

An optical circulator [21] is a three-port device that has a cyclic (non-reciprocal) connectivity. Referring to Figure B.1a), the required functionality is as follows. The polarized output from the laser transmitter (Tx) is the input into port 1 of the circulator and is directed to port 2 which is connected to the optical fiber. Because standard optical fiber does not maintain the polarization state [43], the output lightwave signal for a bidirectional link from the fiber at the input to port 2 (right side of the figure) has a random polarization state. This light is directed to port 3, which is the input to the receiver (Rx).

Figure B.1b) shows a schematic representation of an integrated form of an optical circular. This circulator manipulates two polarization states (indicated as "s" and "p") of the lightwave signal and has three elements. The first element is a set of polarizing beam splitters (PBS) shown as the four gray lines at 45^o in Fig. B.1b). These devices transmit one state of polarization ("p") and reflect the orthogonal state of polarization ("s"). The second device is a Faraday rotator (FR). This device is based on a magneto-optic material, which for the device under consideration, has the property that the polarization plane rotates by $\pm 45^o$ with the sign of the rotation depending on the direction of light propagation through the rotator. This is a non-reciprocal device. The third device is a birefringent wave plate (HWP) that is used to rotate the polarization plane by 45^o. This is a reciprocal device. When the polarized light (red arrow) from the laser transmitter (Tx) enters the circulator from the right side, the polarization plane rotates by -45^o when passing through the Faraday rotator and rotates by 45^o when passing through the wave plate. These two polarization rotations cancel so that the state of polarization remains the same as indicated by the color of the red arrow not changing when traversing from port 1 to port 2 of the circulator.

Now consider the input into port 2 of the circulator. When the unpolarized light (blue and red arrows) exiting the fiber enters port 2 of the circulator on the left, the two polarizations are separated by the polarization beamsplitter. Each of two polarizations is rotated by 45^o when passing through the wave plate and then rotated by an additional 45^o when passing through the Faraday rotator with the sign of the rotation changing because the lightwave is propagating in the opposite direction through the (non-reciprocal) rotator. The net result is that the polarization plane for each of the two polarizations is rotated by 90^o as indicated by the color of the arrows in Fig. B.1b) changing from blue to red or red to blue. Another polarization beamsplitter in the upper right corner of the figure combines these two polarizations at the output of port 3, which is the input to the optical receiver (Rx).

Because of the numerous non-standard requirements for the transceivers, the initial implementations of the circulator were external to the optical transceivers. Other versions of the transceiver integrated the circulator into the transceiver module for further performance, size, and cost reduction at the expense of the ability to reuse across different generation transceivers. Fig. C.1 shows an example implementation of circulator integration within a single pluggable transceiver module.

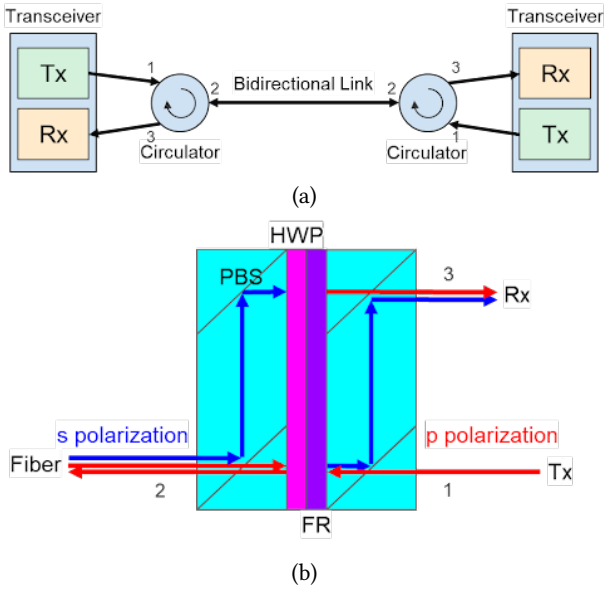

Figure B.1: a) The optical circulator is a three-port non-reciprocal device that has a cyclic connectivity. Input into port 1 is directed to port 2, input into port 2 is directed to port 3. The circulator thus converts a traditional duplex optical transceiver that uses two fiber strands into a bidirectional one that uses one strand, saving 50% of the OCS ports required for operation. b) Example integrated circulator implementation. Lines indicate the directions of lightwave polarization with the blue lines showing s-polarization and red lines indicate p-polarization. PBS - polarizing beam splitter, FR - faraday rotator, HWP - half wave plate. Numbers indicate corresponding port numbers matching those shown in (a)

C TRANSCEIVER AND OCS TECHNOLOGIES

This Appendix presents background material on transceiver and optical switching technologies.

C.1 Transceiver Technology

The section provides background on the existing standards-based transceiver technology roadmap and highlights the key differences between this roadmap and the transceiver technology developed for use in our lightwave fabrics. Figure 8 shows the WDM single mode transceiver roadmap developed over the past decade [34–36]. The initial development of the transceivers developed for the DCN and ML lightwave fabric use cases diverged away from this roadmap in three key areas, which are shown diagrammatically in Fig. C.1.

The use of an integrated circulator enabled the bidirectional links. The use of externally modulated lasers (EMLs) was critical for mitigating multi-path interference (MPI) effects enhanced by bidirectional communication. The use of custom DSP blocks [10]

Technology	Relative Cost*	Port Count	Switching Time	Insertion Loss (dB)**	Driving Voltage (V)	Latching
MEMS [7, 48]	Medium	320×320	milliseconds	<3	≈100	No
Robotic [23]	Medium	1008×1008	minutes***	< 1	NA	Yes
Piezo [46]	High	576×576	milliseconds	<2.5	≈ 10	No
Guided Wave [55]	Low	16×16	nanoseconds	<6	≈ 1	No
Wavelength[65]	TBD	100×100	nanoseconds	<6	0	Yes

* Based on the scale indicated ** Includes connector losses *** Per connection

Table C.1: Cost, scale, performance, and reliability/availability comparison of various OCS technologies. Latching refers to the ability of the OCS to maintain its switch state after a power failure.

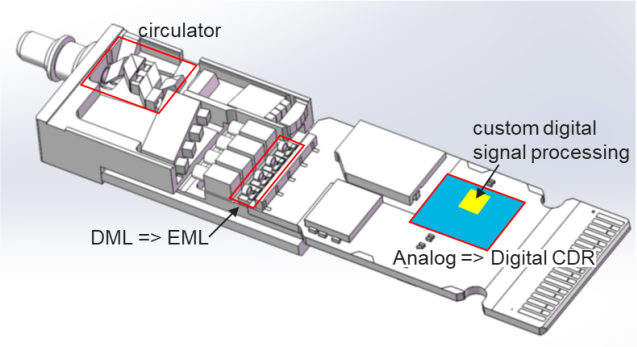

Figure C.1: Bidirectional CWDM4 optical transceiver showing three key differences compared to standards-based transceivers: the integrated circulator, the change from a directly modulated laser (DML) to an externally modulated laser (EML), and the custom DSP blocks used to mitigate interference and provide advanced error correction.

provided a more robust, scalable solution by relaxing the requirements on optical and analog electrical components. DSP blocks were also developed to mitigate interference impairments inherent in bidirectional links [15, 58, 60]. This also includes the development of soft forward error correction (sFEC) techniques in order to support the higher link budgets needed. We note that our work in 100 GbE WDM transceivers eventually led to the creation of the CWDM4 MSA [2]. A variant of the advanced FEC has been adopted by the 200 Gb/s PAM4 IEEE Standards Group (IEEE802.3dj) [44].

C.2 OCS Technology

This section provides background material on optical circuit switching technologies. Table C.1 compares cost, scale, performance, and reliability/availability of various OCS technologies that could be used for large-scale applications. Systems employing piezo-electric actuation, robotics to mechanically reconfigure a patch panel, and MEMS are among those previously achieving some limited commercial adoption. In terms of scaling to the large number of port counts required for scale-out applications at acceptable costs, MEMS-based systems had demonstrated the most promise, with the realization of systems achieving greater than 1000×1000 interconnectivity [48]. The robotics configured OCS, although able to scale to larger port counts while supporting any fiber type (single and multi mode fibers), suffers from slow switching speeds that are further compounded by the need to serialize switching of connections. Piezo-based systems can have higher costs due to assembly complexity, but can also scale to large port counts with commercial systems available with 576 ports [46]. Guided wave switching, such as PLZT-based switching, has limited scale with high losses, although offering the lowest costs and size due to integration.

Wavelength-switching-based schemes have been extensively investigated by the research community, primarily for faster optical switching (optical packet switching (OPS)/optical burst switching (OBS)) within core telecommunication networks. These systems use a combination of tunable lasers, array waveguide devices, and tunable filters. However, even for slower switching applications, wavelength switching lacks future proofing, as the channel spacing and width of the arrayed waveguide grating (AWG) limits the link speed and wavelength plan.