Accelerating DML Training in Optical Rackless DC with P4-based In-Network Computing

(Invited Paper)

Weichi Wu, Ag Pang, Youyou Ai, Qiqi Zhu, Zuqing Zhu†
School of Information Science and Technology, University of Science and Technology of China, Hefei, China
†Email: {zqzhu}@ieee.org

***Abstract*—This work describes a novel data-center network (DCN) architecture (namely, P4-ORDC) to explore the mutual benefits of optical rackless data center (ORDC) and in-network computing (INC) for accelerating DML training jobs effectively. We elaborate on the operation principle, system design, and implementation of P4-ORDC, and prototype a small-scale network system to demonstrate its effectiveness experimentally. The experimental results indicate that P4-ORDC reduces job completion time (JCT) over existing benchmarks.**

***Index Terms*—Rackless data-center, Distributed machine learning, P4, In-network computing, Collective communications.**

I. Introduction

Despite the fast development of the Internet over the past decade [1–10], the recent popularity of artificial intelligence (AI) has led to fast development of large language model (LLMs), whose training can hardly be handled by single graphic processing units (GPUs) [11] and thus brings in new networking challenges. Hence, distributed machine learning (DML) jobs (*e.g.*, distributed training of LLMs) became the fastest-growing applications in data-center networks (DCNs) [12]. Other than legacy DCN applications, DML jobs rely more on collective communications, each instance of which organizes a cluster of processes to take part in synchronized computing and communication phases iteratively [13]. In a computing phase, each process finishes its share of training task, and thus very little traffic is induced. On the other hand, the processes exchange their training results in a communication phase, generating huge volumes of traffic in skewed patterns. Then, the rising of DML jobs significantly changes traffic patterns in DCNs, pushing harder for network reconfigurability, throughput and energy efficiency.

Introducing optical circuit switching (OCS) into DCN can potentially address the challenges DML jobs brought to D-CNs [14–27]. This is because other than traffic engineering (TE), OCS-based DCN (ODCN) enables topology engineering (TPE) to allow for extra flexibility [28]. As described in [29–34], an ODCN can interconnect top-of-rack (ToR) switches using OCS switches (*e.g.*, optical cross-connects (OXCs)), with which TPE steers inter-rack connectivity to adapt to dynamic and skewed traffic, offering another level of network reconfigurability. However, such an ODCN configuration might not fully resolve the major network bottleneck caused by DML jobs at the rack-level. Specifically, the topological rack boundaries in a DC can still fragment DML jobs across racks,

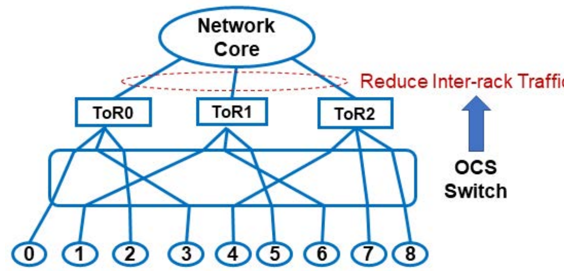

Fig. 1. Example on optical rackless data-center (ORDC) pod.

while inter-rack network bottleneck will restrict the throughput of their communication phases and eventually prolong their training time. This dilemma can be solved by considering the optical rackless data-center (ORDC) [35] that removes the topological rack boundaries in DCs by putting an OCS switch in between servers and ToR switches to adapt to the dynamic traffic patterns generated by DML jobs. An ORDC pod example is shown in Fig. 1, where an OCS switch connects 8 servers in the pod to three ToR switches, facilitating various rack configuration to adapt to application demands.

Although the existing work in [35] has verified the advantages of ORDC, the training of DML jobs can be performed more efficiently if we can integrate P4-based in-network computing (INC) [36] into ORDC. The advances on programmable data plane (PDP) [37] have facilitated aggregation and arithmetic operations on packet fields at line-rate directly in PDP switches. Hence, people can leverage INC to offload certain computing tasks in DML training from servers to PDP-based ToR switches (namely, P-ToR switches), achieving effective acceleration [38]. Taking the DML training in the parameter-server framework as an example, Fig. 2 illustrates the INC-based acceleration. The each iteration of the training includes four phases: 1) *map* (the master broadcasts data to workers), 2) *local compute* by the workers, 3) *upload* (the workers incast their results to the master), and 4) *reduce* (the master aggregates workers' results). As the aggregation operation in the fourth phase can be realized with simple arithmetic operations, it can be offloaded to a P-ToR switch with INC such that the fourth phase is performed simultaneously with the third one and accelerates the DML training effectively. In the meantime, as the P-ToR switch only needs to send INC-aggregated results to the master, it essentially absorbs half of

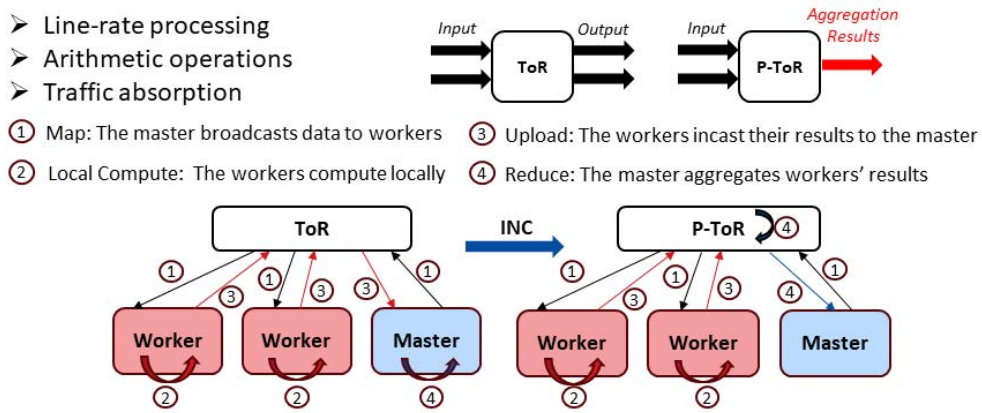

Fig. 2. Example in INC-based computing acceleration with PDP switches.

the worker-to-master traffic in the third phase.

To this end, we propose to integrate ORDC with INC for a novel pod-level DC architecture (namely, P4-ORDC), to explore their mutual benefits as follows. The reconfigurability enabled by ORDC enables time sharing of INC resources (*i.e.*, registers and arithmetic logic units (ALUs)) on P-ToR switches, improving their utilization over time to accelerate offloadable computing phases in DML training better. The traffic absorption achieved by INC relieves network bottlenecks in each ORDC pod, shortening communication phases of DML training. In this work, we elaborate on the design and implementation of P4-ORDC, which partially replaces ToR switches in an ORDC pod with P-ToR switches, program P4-based packet processing pipelines in them to enable INC, and develop a control plane to explore the mutual benefits of ORDC and INC. In order to demonstrate its performance, we prototype P4-ORDC with commercial products and build a small-scale ORDC pod. Experimental results show that our proposed P4-ORDC significantly outperforms existing benchmarks (ORDC without INC [35] and optically-interconnected P-ToR switches [28]) on average job completion time (JCT).

The rest of this paper is organized as follows. Section II explains the operation principle of P4-ORDC and the system design, while the experimental results are discussed in Section III. Finally, Section IV summarizes the paper.

II. Operation Principle and System Design

The data plane of P4-ORDC is shown in Fig. 3, which generally follows the architecture of the ORDC pod in [35], except that we replace certain ToR switches in it with P-ToR switches and implement a packet encoder in each server to assist INC operations in the P-ToR switches. We leverage the data plane development kit (DPDK) to develop the packet encoder for ensuring high-speed packet processing and transmitting. Specifically, we program the DPDK-based packet encoder to 1) intercept all the workers-to-master packets in the third phase (*upload*) and 2) encode them in a customized packet format for INC. Meanwhile, we also design and implement the P4 pipeline in the P-ToR switch to realize INC, which makes the third and fourth phases in Fig. 2 happen concurrently.

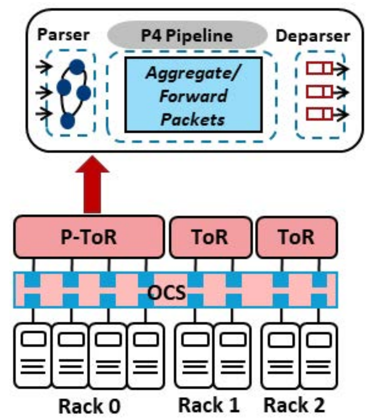

Fig. 3. Data plane design of P4-ORDC.

Fig. 4 illustrates the control plane of P4-ORDC, which is crucial to explore the mutual benefits of ORDC and INC to accelerate DML training. As for the ToR switches, servers, and OCS, we allocate the ToR controller, job handler, and OCS controller to respectively manage them, through corresponding south-bound interfaces. The ToR controller collects the statistics of traffic through ToR switches (both traditional ToR and P-ToR switches) to update the traffic engineering database (TED). The servers in the P4-ORDC pod register their DML jobs to the job handler, which checks the JobID database to assign a unique *JobID* to each job and forward the jobs' information (durations of their computing phases, data sizes of their communication phases, *etc.*) to the TED. The job handler also interacts with the INC scheduler to assign INC resources in P-ToR switches to offloadable computing phases. The TED analyzes the information about traffic and jobs in the data plane to abstract a 3-dimensional (3D) traffic matrix that represents the data transfers between server pairs over a series of future time periods. Then, as shown in the top subplot of Fig. 4, the topology engineering (TPE) module obtains a sequence of OCS configurations, each of which has been optimized to group servers into racks such that the DML training during the corresponding future period can be accelerated the most. Finally, the OCS configurations will be applied to the OCS as planned by the OCS controller, and the TED will calculate the routing schemes for traffic based on each OCS configuration and implement them in the ToR switches through the ToR controller.

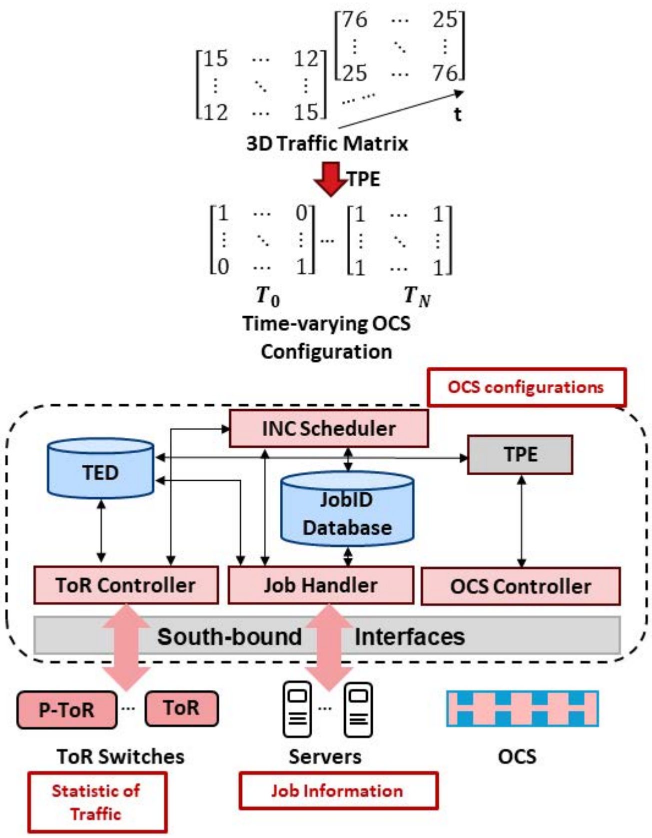

Fig. 4. Control plane design of P4-ORDC.

III. Experimental Demonstration

We build a small-scale P4-ORDC pod with 8 servers, three ToR switches, and a 32×32 optical cross-connect (OXC), and the line-rate of the links to connect them is 10 Gbps. One ToR switch is a P-ToR switch based on Tofino ASICs. The physical setup of the P4-ORDC pod is the same as that in Fig. 3, where the P-ToR switch connects to the OXC with four ports and each traditional ToR allocates two ports to the OXC. For convenience, we index the servers in Fig. 3 from left to right as *Servers* 0-7, respectively, and for simplicity, we run Hadoop WordCount jobs instead of DML training jobs on them, which have similar collective communication pattern but with much simpler INC procedure. We first conduct a simple experiment to demonstrate the mutual benefits of ORDC and INC on accelerating collective communications illustratively. Specifically, we place two sets of jobs on *Servers* 0-3 and 4-7, respectively, and let P4-ORDC orchestrate the collective communications of the jobs to share INC resources on the P-ToR switch over time. Fig. 5 shows the jobs' traffic that goes through the P-ToR switch, indicating that the control plane schedules the *upload* phases of the two server groups alteratively over time. In this experiment, the mutual benefits of ORDC and INC achieve an average JCT reduction of 66%.

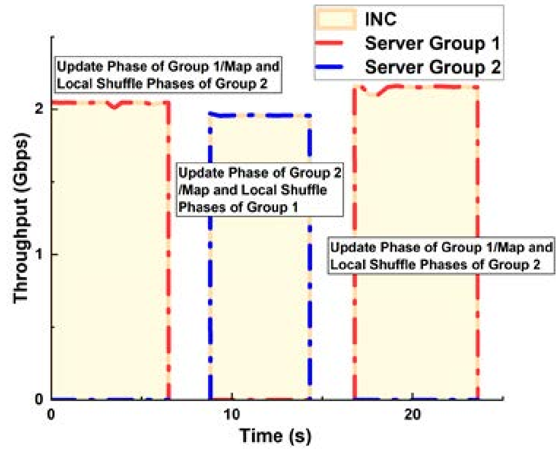

Fig. 5. Results on interleaving jobs with collective communications.

We then conduct experiments to compare P4-ORDC to three existing architectures: 1) FatTree, which removes the OXC, replaces the P-ToR switch with a normal ToR switch, and interconnects the ToR switches with Ethernet switches in a fat-tree, 2) ORDC [35], the traditional ORDC without INC, and 3) P4INC-AOI [28], which interconnects P-ToR and ToR switches from the top with an OXC. For fair comparisons, the numbers of servers, ToR switches and ports on each ToR switch are kept the same in the four architectures. The results on average JCT are plotted in Fig. 6. Due to the lack of network reconfigurability and INC, FatTree provides the longest JCT in all the experimental scenarios, and it is followed by ORDC, justifying the necessity of network reconfigurability in serving collective communications. INC can accelerate collective communications effectively, since P4INC-AOI and P4-ORDC achieve significantly shorter JCTs than the other two, but compared with P4INC-AOI, P4-ORDC further reduces the average JCT by 23.4% to 26.4%, confirming the mutual benefits of ORDC and INC.

IV. Conclusion

In this paper, we discussed P4-ORDC, which is a novel DCN architecture to explore the mutual benefits of ORDC and INC for accelerating DML training jobs with collective communications. We prototyped a small-scale P4-ORDC with off-the-shelf network elements, and our experimental results verified that P4-ORDC can effectively reduce JCT over a few existing benchmark DCN architectures.

Acknowledgments

This work was supported by the NSFC project 62371432.

References

[1] "Cisco Annual Internet Report 2024." [Online]. Available: https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/index.html#~executive-summaries.

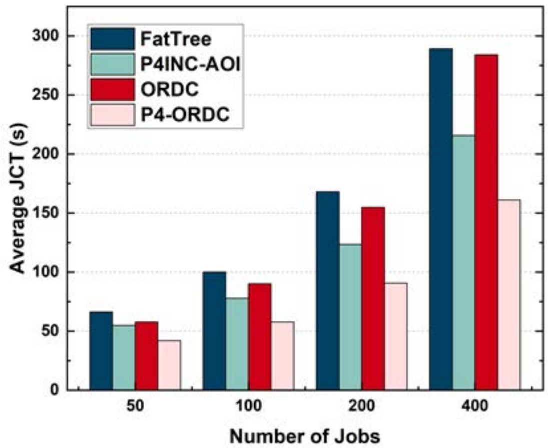

Fig. 6. Results on JCT in different DCN architectures.

[2] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," *J. Lightw. Technol.*, vol. 31, pp. 15–22, Jan. 2013.
[3] L. Gong *et al.*, "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," *J. Opt. Commun. Netw.*, vol. 5, pp. 836–847, Aug. 2013.
[4] Y. Yin *et al.*, "Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 5, no. 10, pp. A100–A106, Oct. 2013.
[5] L. Gong and Z. Zhu, "Virtual optical network embedding (VONE) over elastic optical networks," *J. Lightw. Technol.*, vol. 32, pp. 450–460, Feb. 2014.
[6] P. Lu *et al.*, "Highly-efficient data migration and backup for Big Data applications in elastic optical inter-datacenter networks," *IEEE Netw.*, vol. 29, pp. 36–42, Sept./Oct. 2015.
[7] W. Lu, Z. Zhu, and B. Mukherjee, "On hybrid IR and AR service provisioning in elastic optical networks," *J. Lightw. Technol.*, vol. 33, pp. 4659–4669, Nov. 2015.
[8] J. Liu *et al.*, "On dynamic service function chain deployment and readjustment," *IEEE Trans. Netw. Serv. Manag.*, vol. 14, pp. 543–553, Sept. 2017.
[9] P. Lu and Z. Zhu, "Data-oriented task scheduling in fixed- and flexible-grid multilayer inter-DC optical networks: A comparison study," *J. Lightw. Technol.*, vol. 35, pp. 5335–5346, Dec. 2017.
[10] S. Li *et al.*, "Protocol oblivious forwarding (POF): Software-defined networking with enhanced programmability," *IEEE Netw.*, vol. 31, pp. 58–66, Mar./Apr. 2017.
[11] D. Patel and G. Wong. (2023, Jul.) GPT-4 architecture, infrastructure, training dataset, costs, vision, MoE. [Online]. Available: https://www.semianalysis.com/p/gpt-4-architecture-infrastructure.
[12] J. Verbraeken *et al.*, "A survey on distributed machine learning," *ACM Comput. Surv.*, vol. 53, pp. 1–33, Mar. 2020.
[13] C. Cameron, S. Lantz, and B. Barker. What is collective communication? [Online]. Available: https://cvw.cac.cornell.edu/mpicc/intro/index.
[14] H. Liu *et al.*, "Lightwave Fabrics: At-scale optical circuit switching for datacenter and machine learning systems," in *Proc. of ACM SIGCOMM 2023*, pp. 499–515, Sept. 2023.
[15] H. Fang *et al.*, "Predictive analytics based knowledge-defined orchestration in a hybrid optical/electrical datacenter network testbed," *J. Lightw. Technol.*, vol. 37, pp. 4921–4934, Oct. 2019.
[16] Z. Zhu, W. Lu, L. Liang, and B. Kong, "Predictive analytics in hybrid optical/electrical DC networks," in *Proc. of OFC 2019*, pp. 1–3, Mar. 2019.
[17] W. Lu *et al.*, "AI-assisted knowledge-defined network orchestration for energy-efficient data center networks," *IEEE Commun. Mag.*, vol. 58, pp. 86–92, Jan. 2020.
[18] S. Zhao and Z. Zhu, "Network service reconfiguration in hybrid optical/electrical datacenter networks," in *Proc. of ONDM 2020*, pp. 1–6, May 2020.
[19] Q. Li *et al.*, "Scalable knowledge-defined orchestration for hybrid optical/electrical datacenter networks," *J. Opt. Commun. Netw.*, vol. 12, pp. A113–A122, Feb. 2020.
[20] S. Zhao and Z. Zhu, "On virtual network reconfiguration in hybrid optical/electrical datacenter networks," *J. Lightw. Technol.*, vol. 38, pp. 6424–6436, Dec. 2020.
[21] X. Pan *et al.*, "Scheduling virtual network reconfigurations in parallel in hybrid optical/electrical datacenter networks," *J. Lightw. Technol.*, vol. 39, pp. 5371–5382, Sept. 2021.
[22] S. Zhao, X. Pan, and Z. Zhu, "On the parallel reconfiguration of virtual networks in hybrid optical/electrical datacenter networks," in *Proc. of GLOBECOM 2020*, pp. 1–6, Dec. 2020.
[23] H. Yang and Z. Zhu, "Application-aware configuration of all-optical interconnects in Hyper-FleX-LION," in *Proc. of OGC 2022*, pp. 1–4, Sept. 2022.
[24] J. Peng *et al.*, "Offloading NFV orchestration to ToR switches: How to leverage PDP to realize agile service function chaining in HOE-DCNs," in *Proc. of ICC 2021*, pp. 1–6, Jun. 2021.
[25] H. Yang *et al.*, "On the bilevel optimization for remapping virtual networks in an HOE-DCN," *IEEE Trans. Netw. Serv. Manag.*, vol. 19, pp. 1274–1286, Jun. 2022.
[26] H. Yang, Z. Zhu, R. Proietti, and B. Yoo, "Which can accelerate distributed machine learning faster: Hybrid optical/electrical or optical reconfigurable DCN?" in *Proc. of OFC 2022*, pp. 1–3, Mar. 2022.
[27] Q. Lv, Z. Ma, and Z. Zhu, "Experimental demonstration of hitless OCS-based DCN reconfiguration to steer multi-class traffic," in *Proc. of ICOCN 2024*, pp. 1–3, Jul. 2024.
[28] X. Xie, H. Yang, and Z. Zhu, "P4INC-AOI: When in-network computing meets all-optical interconnect for adaptive and low-latency optical DCN," in *Proc. of OFC 2023*, pp. 1–3, Mar. 2023.
[29] H. Yang and Z. Zhu, "Acceleration of coflows with all-optical interconnects in Hyper-FleX-LION," *J. Opt. Commun. Netw.*, vol. 14, pp. 805–814, Oct. 2022.
[30] M. Zhang *et al.*, "Gemini: Practical reconfigurable datacenter networks with topology and traffic engineering," *arXiv preprint arXiv:2110.08374*, Oct. 2021.
[31] H. Yang and Z. Zhu, "Traffic-aware configuration of all-optical data center networks based on Hyper-FleX-LION," *IEEE/ACM Trans. Netw.*, vol. 32, pp. 2675–2688, Jun. 2024.
[32] Q. Lv *et al.*, "On the TPE design to efficiently accelerate hitless reconfiguration of OCS-based DCNs," *IEEE J. Sel. Areas Commun., in Press*, pp. 1–13, 2025.
[33] X. Xie, B. Tang, X. Chen, and Z. Zhu, "P4INC-AOI: All-optical interconnect empowered by in-network computing for DML workloads," *IEEE/ACM Trans. Netw., in Press*, pp. 1–16, 2025.
[34] X. Chen *et al.*, "DRL-TPE: Learning to optimize TPE of optical interconnects to accelerate hitless reconfiguration of OCS-based DCNs," in *Proc. of OFC 2025*, pp. 1–3, Mar. 2025.
[35] W. Wang *et al.*, "RDC: Energy-efficient data center network congestion relief with topological reconfigurability at the edge," in *Proc. of NSDI 2022*, pp. 1267–1288, Apr. 2022.
[36] X. Dong *et al.*, "On scheduling DML jobs in all-optical DCNs with in-network computing," in *Proc. of GLOBECOM 2024*, pp. 1–6, Dec. 2024.
[37] P. Bosshart *et al.*, "P4: Programming protocol-independent packet processors," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, pp. 87–95, Jul. 2014.
[38] W. Wu *et al.*, "Accelerating collective communications with mutual benefits of optical rackless DC and in-network computing," in *Proc. of OFC 2025*, pp. 1–3, Mar. 2025.