Performance Comparison of Distributed DNN Training on Optical versus Electrical Interconnect Systems

Fei Dai[1,2(✉)], Yawen Chen[1], Zhiyi Huang[1], Haibo Zhang[1], and Hui Tian[3]

[1] School of Computing, University of Otago, Dunedin, New Zealand
[2] School of Computing, Eastern Institute of Technology | Te Pūkenga, Napier, Hawke's Bay, New Zealand
[3] School of Information and Communication Technology, Griffith University, Brisbane, Australia
travis.dai@otago.ac.nz

Abstract. Parallel and distributed Deep Neural Network (DNN) training have become integral in data centers, significantly reducing DNN training time. The interconnection type among nodes and the chosen all-reduce algorithm critically impact this speed-up. This paper examines the efficiency differences in distributed DNN training across optical and electrical interconnect systems using various all-reduce algorithms. We first explore the Ring and Recursive Doubling (RD) all-reduce algorithms in both systems, followed by formulating a communication cost model for these algorithms. Performance comparison is then carried out via extensive experiments. Our results reveal that, in 1024-node systems, the Ring algorithm outperforms the RD algorithm in optical and electrical interconnects when data transfer exceeds 64MB and 1024MB, respectively. We also find that both Ring and RD algorithms in optical interconnect systems reduce average communication time by around 75% compared to electrical interconnect systems across four different DNNs. Interestingly, the communication time of the RD algorithm, but not the Ring algorithm, reduces as the number of wavelengths increase in optical interconnects. These findings provide valuable insights into DNN training optimization across various interconnect systems and lay the groundwork for future related research.

Keywords: Performance comparison · Distributed DNN training · Electrical interconnect · Optical interconnect · Performance comparison · Parallel computation.

1 Introduction

Deep Neural Networks (DNNs), fundamental algorithms in Deep Learning (DL), have found widespread applications in domains such as image classification, language translation, and speech recognition [1]. However, training DNN models

with large datasets can be a painstakingly long process, often taking days or even weeks, which is not practical in most cases. As a response, numerous parallel and distributed DNN training methods have emerged. Among these, data parallelism is widely adopted: each worker trains the DNN using its local dataset and iteratively exchanges model parameters (e.g., gradients) with other workers. Stochastic Gradient Descent (SGD), the leading method for DNN training, invokes intense communications for the all-reduce operation in distributed DL training [14]. However, with an escalating number of workers, network communication traffic balloons, necessitating higher bandwidth.

Electrical interconnect systems for distributed DNN training are advancing rapidly. Yet, their bandwidth improvements lag behind the progression of computing units [15]. High communication overheads from factors like crosstalk, dielectric loss, and switching noise render electrical interconnect a critical obstacle to enhancing overall system performance [16]. When communication overhead outweighs the gains from parallel computation, training performance deteriorates. Optical interconnect emerges as a promising alternative [23], offering high bandwidth, low power costs, and minimal latency. Leveraging wavelength division multiplexing (WDM), optical interconnect allows data transmission through a waveguide using different wavelengths. These benefits position optical interconnects as efficient performers in intensive device-to-device communications and can thus accelerate data-parallel distributed DNN training.

Nonetheless, optical interconnect also carries overheads such as optical-to-electrical (OE) / electrical-to-optical (EO) conversion costs, insertion loss from light transmission through the waveguide, and configuration delay of the micro-ring resonator (MRR). These factors, along with communication patterns governed by the all-reduce algorithm, influence the communication time in data-parallel distributed DNN training. To date, no studies have compared the communication performance of all-reduce algorithms between optical and electrical interconnect systems in data-parallel distributed DNN training.

Thus, this paper investigates the comparison of all-reduce algorithms for distributed DNN training efficiency on optical and electrical interconnect systems under varying configurations. We aim to answer the following research questions: *1) How does the performance of the Ring and Recursive Doubling (RD) all-reduce algorithms differ between optical and electrical interconnect systems with varying sizes of transferred data? 2) What improvements can optical interconnect systems bring to distributed DNN training compared with electrical systems? 3) Are existing all-reduce algorithms utilizing optical resources in the optical interconnect system optimally for distributed DNN training?* The key contributions of this paper are summarized as follows:

1. We first analyze Ring and RD all-reduce algorithms and summarize their communication step and the amount of communication traffic in each step. Then, we formulate the communication costs of these two algorithms on the optical and electrical interconnect systems for data-parallel distributed DNN training.

2. We conduct extensive simulations to compare the communication performance of those two all-reduce algorithms between the optical interconnect system and the electrical interconnect system using two experiment settings. Our results indicate: a) The Ring all-reduce algorithm outperforms the RD algorithm when the transferred data is more than 64MB and 1024MB respectively in the 1024 nodes optical and electrical interconnect systems. b) The optical interconnect system demonstrates superior scalability compared to its electrical counterpart. Additionally, the Ring and RD all-reduce algorithms on the optical interconnect system reduce the average communication time by 74.74% and 75.35% for four DNNs compared with the electrical interconnect system. c) In the optical interconnect system, the Ring all-reduce algorithm's communication time is not impacted by the number of wavelengths. In contrast, the RD algorithm's communication time decreases with an increasing number of wavelengths.

The remaining part of the paper proceeds as follows: Section 2 introduces the related works. Section 3 provides the necessary background on optical and electrical interconnect system and data-parallel distributed DNN training. Section 4 first illustrates the all-reduce algorithms, then presents the formulation of communication costs for the two all-reduce algorithms in optical and electrical interconnect systems. Section 5 outlines the experimental setup and compares the performance of the two all-reduce algorithms across optical and electrical interconnect systems. Finally, Section 6 concludes the paper.

2 Related work

An immense amount of research has been conducted to expedite distributed DNN training in electrical interconnect systems. For instance, Plink [2] employs two-level hierarchical aggregations to balance traffic and curtail communication delay. Blink [3], designed specifically for the DGX-2 GPU, uses multiple directed spanning trees to optimize link utilization. Poseidon, a robust communication framework for distributed DL on GPUs, has been devised to address the difficulties associated with scaling out due to the necessity for regular network synchronization [19]. Other studies have focused on enhancing all-reduce performance by reducing network contention [4], bolstering ring all-reduce performance for large-scale clusters [5], and decomposing all-reduce on heterogeneous network hierarchies [6]. A growing trend also suggests integrating optical interconnect systems for superior performance. For instance, a hybrid electrical-optical switch architecture proposed in [7] achieved a 10% improvement in training time. SipML, another work utilizing optical interconnect for distributed machine learning training, showed a 1.3 - 9.1$\times$ improvement in DNN model training time through simulations [8]. RAMP is an all-optical network architecture that greatly improves energy efficiency and training time for recommendation systems such as Megatron and DLRM [21].

Certain studies have compared the performance between optical and electrical interconnect systems. The study in [9] using the FFT benchmark found

over 70% gain in speedup and efficiency in optical over electrical systems. Another study [10] compared power dissipation, data rate, and loss between the two systems, and simulations revealed higher bandwidth, lower channel loss, and reduced power consumption in optical interconnects. The study referenced in [20] compared the performance of multi-layer perceptron training between electrical and optical network-on-chips. Experimental results demonstrate that the MLP training time for ONoC surpasses ENoC in both training duration and energy efficiency. However, these studies did not focus on comparisons specific to distributed DNN training scenarios. To the best of our knowledge, this paper is the first to compare the performance of optical and electrical interconnect systems in the context of data-parallel distributed DNN training using different all-reduce algorithms.

3 Background

3.1 Optical & Electrical Interconnect Systems

This section first elucidates the major distinctions, benefits, and drawbacks of optical and electrical interconnect systems. It then discusses the particular architectures of these systems used in our study. The fundamental difference between optical and electrical interconnect systems lies in the transmission technology they employ to facilitate communication between hosts or nodes. In electrical interconnect systems, all nodes are connected via electronic interconnect in a multilevel electronic hierarchy, such as a fat tree. Conversely, optical interconnect systems utilize arrays of optical switches or Sip interfaces to interconnect nodes [8]. Communication in electrical interconnect systems transpires through hierarchical electrical switches. Electrical packets from the source traverse through electrical links and switches to reach the destination. In contrast, transmission among nodes in optical interconnect systems leverages different wavelengths for parallel communication via waveguides using WDM technology. The benefits of optical communication include low transmission delay ($10ms$ for optical switch, $25\mu s$ for MRRs Sip interface), low power consumption (essentially independent of distance), and high bandwidth (up to 40 Gb/s per wavelength, with 64 wavelengths per waveguide) [8]. However, one significant disadvantage of optical interconnect systems is the need for numerous optical components, thereby incurring configuration overhead. Compared to optical interconnect systems, electrical interconnect systems offer lower bandwidth (25 Gb/s) and require a much more complex structure for switch interconnect as the cluster scales.

Fig. 1 and Fig. 2 present an overview of the electrical and optical interconnect systems respectively, as used in distributed DNN training. The architecture of the electrical interconnect comprises two levels, where hosts directly connect to level 1 electrical switches. These switches' uplinks connect to level 2 electrical switches, while their downlinks connect to hosts. Each electrical switch features 36 ports. Fig. 2 illustrates the architecture of the optical interconnect system, which aligns with the system proposed in [8] and is based on the Sip interface.

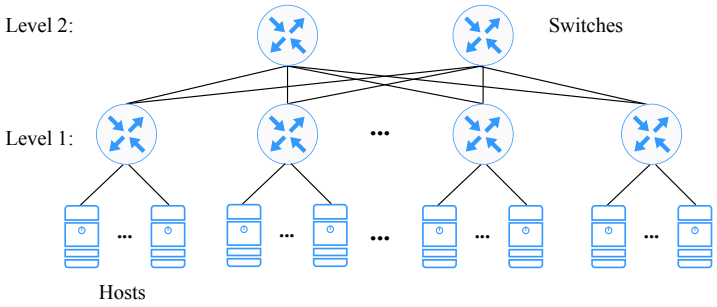

Fig. 1. Electrical interconnect architecture.

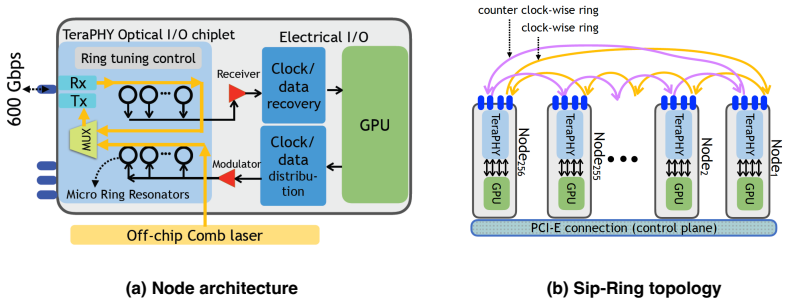

Fig. 2. Optical interconnect architecture: (a) Node architecture, (b) Sip-Ring topology [8].

Fig. 2 (a) shows the composition of nodes leveraging TeraPHY silicon photonics technology. Each node incorporates four optical interfaces, each containing 64 micro-ring resonators (MRRs) to select and forward any subset of 64 wavelengths. On the transmit side (Tx), an off-chip comb laser generates light steered into the node via a fiber coupler towards an array of MRRs. These modulate the accelerator's transmitting data at 40 Gbps per wavelength. Conversely, the receive side (Rx) features a second array of MRRs that select wavelengths targeted to the accelerator and pass through the remaining wavelengths. Fig. 2 (b) depicts nodes interconnected with their adjacent nodes in a ring topology (only two rings for two directions are displayed for clarity). In the data plane, traffic transmits across four single-mode fiber rings, two in each direction (clockwise and counter-clockwise). Routing and wavelength assignment (RWA) configuration is performed in the control plane, where wavelengths can be dynamically placed around the fiber ring. We assume that each host in both electrical and optical interconnect systems uses a GPU as its computing device, and all GPUs across hosts are homogeneous. The details of system parameters are described in Section 5.

The primary motivations for comparing these two interconnect architectures are twofold: (1) to evaluate the distributed DNN training performance of existing all-reduce algorithms on state-of-the-art optical interconnections composed of MRR optical interfaces; and (2) to determine the performance improvement an optical interconnect system can offer compared to a traditional multi-tier, fat-tree topology electrical interconnect system.

3.2 Distributed Data-parallel Training of DNN

A DNN typically contains L layers ($L \geq 2$), among which common types are fully connected, convolutional, attention, and others. This paper mainly focuses on DNNs with fully connected and convolutional layers. Given that each DNN layer can be a fully connected or convolutional layer, the neurons in layer ℓ receive vectors or tensors as input. We first elucidate the scenario for a fully connected layer, then extend it to the convolutional layer. The DNN training process encompasses forward propagation (FP) and backward propagation (BP). In forward propagation, we denote the output vector/tensor in layer ℓ (or the input vector/tensor of layer $\ell - 1$) as $Z^{(\ell)}$, the weight matrix at layer ℓ as $W^{(\ell)}$, and the bias vector/tensor in layer ℓ as $B^{(\ell)}$. Thus, the forward propagation of DNN training with n_ℓ neurons at layer ℓ can be defined as:

$$Z^{(\ell)} = \phi(W^{(\ell)} Z^{(\ell-1)} + B^{(\ell)}), \ell = 2, 3, ..., L. \tag{1}$$

where $\phi(*)$ is the activation function.

During backward propagation, we define the error vector/tensor in layer ℓ as E_ℓ. We can then express this as:

$$E^{(\ell)} = (E^{(\ell+1)}(W^{(\ell)})^T)\phi'(Z^{(\ell)}), \tag{2}$$

where $\phi'(*)$ is the derivative function of $\phi(*)$.

Using the error vector/tensor, the gradient of the weight $\Delta W^{(\ell)}$ can be calculated as:

$$\Delta W^{(\ell)} = (Z^{(\ell)})^T E^{(\ell+1)}. \tag{3}$$

Upon obtaining the gradient, weights are updated as follows:

$$W^{(\ell)} = W^{(\ell)} - \sigma \Delta W^{(\ell)}, \tag{4}$$

where σ is the learning rate.

Given that a convolutional layer can also be transformed into a matrix-matrix multiplication by the im2col function [11], albeit with different matrix operand dimensions, the training of the convolutional layer can be represented by the above equations.

In data-parallel distributed DNN training, training is parallelized by partitioning the input data across a batch size of b. The whole training model is replicated, such that each worker maintains a local copy but processes different training batches. In the weight update process of BP, all-reduce algorithms synchronize the gradients among the workers. Assuming n nodes in the cluster system, the update process in data-parallel distributed DNN training becomes:

$$\begin{aligned} W^{(\ell)} &= W^{(\ell)} - \sigma \Delta W^{(\ell)} \\ &= W^{(\ell)} - \frac{\sigma(\Delta W_1^{(\ell)} + \Delta W_2^{(\ell)} + ... + \Delta W_n^{(\ell)})}{n}. \end{aligned} \tag{5}$$

4 Performance model

4.1 All-reduce algorithm

During the data-parallel distributed DNN training process (specifically, the weight update process of BP), all nodes in the interconnect systems need to receive data from all other nodes using all-reduce algorithms. Consequently, the communication pattern and communication traffic are dependent on the specific all-reduce algorithms used. Given the distinct characteristics of electrical and optical interconnect systems, the communication steps and sizes of transferred data can vary across different all-reduce algorithms, leading to different training times. Hence, we consider two well-known communication schemes for performing all-reduce operations in a system with n nodes and d as the size of transferred data to be reduced:

– **Ring all-reduce**: The communication in Ring all-reduce consists of two stages - reduce and broadcast stage. Ring all-reduce first reduce-scatters data of size $\frac{d}{n}$ on all nodes in $n-1$ steps during the reduce stage. Then, all nodes conduct all-gather operations to collect reduced data of size $\frac{d}{n}$ from other nodes until every node obtains reduced data of size d. This process also takes $n-1$ steps. Therefore, the total number of steps required by the Ring all-reduce is $2(n-1)$ steps. Fig. 3 (a) illustrates how all-reduce operations are conducted by the Ring all-reduce algorithm in a 4-node optical interconnect

system. The Ring all-reduce communications comprise two stages, each with three steps. Since reduction operations are conducted every step during the reduce stage and all-gather operations are conducted every step during the broadcast stage, the transferred data in each communication step is $\frac{d}{4}$. During each communication step, each node sends data to its adjacent node using one wavelength.

– **Recursive Doubling (RD) all-reduce**: RD with pairwise exchange enables logarithmic scaling of O($log_2 n$). During each communication step, each node in the group sends data of size d to its corresponding peer appropriate for the current step. After nodes receive the data, reduction operations are executed in parallel. The overall all-reduce operation among all nodes is built up by recursively conducting a series of smaller all-reduce operations over orthogonal sub-groups of nodes until the entire target group has been reduced. Fig. 3 (b) demonstrates how the all-reduce communications are conducted by RD in a 4-node optical interconnect system. In the first step, nodes (nodes 1 and 2) in group one and nodes (nodes 3 and 4) in group two exchange data using two wavelengths. In the second step, nodes 1 and 3 form a new group, and nodes 2 and 4 form another group. Then, nodes in the groups exchange the reduced data within the group. According to the communication pattern of RD, the number of wavelengths requirement doubles in every step, and the size of transferred data in each step is d.

4.2 Communication cost

Communication step The communication time in distributed DNN training is influenced by the number of communication steps and the amount of communication traffic in each step, which are determined by the all-reduce algorithm employed in the interconnect systems. Hence, we examine the number of communication steps involved in the Ring and RD all-reduce algorithms, as well as the amount of communication traffic generated in each communication step for both optical and electrical interconnect systems.

In optical interconnect systems, the number of communication steps in the all-reduce algorithm depends on the communication patterns and the available number of wavelengths. In Ring all-reduce, nodes simultaneously send data of size $\frac{d}{n}$ to adjacent nodes, taking $2(n-1)$ steps. For RD all-reduce, nodes send data of size d to their corresponding pair in each step, with multiple wavelengths used concurrently to prevent conflict. When the required number of wavelengths exceeds the available number in the system, additional communication steps are needed. For instance, if the available number of wavelengths is 64, a system with 128 nodes would need $\lceil log_2 128 \rceil = 7$ steps to complete the all-reduce operations, while a system with 1024 nodes would need $2\lceil log_2 1024 \rceil - 7 = 13$ steps to do so. Contrarily, in an electrical interconnect system, which is based on packet switching, the number of communication steps for Ring and RD all-reduce are $2(n-1)$ and $\lceil log_2 n \rceil$, respectively. We summarize the communication steps for the corresponding interconnect systems in Table 1, where the number of available wavelengths in the optical interconnect system is denoted as w.

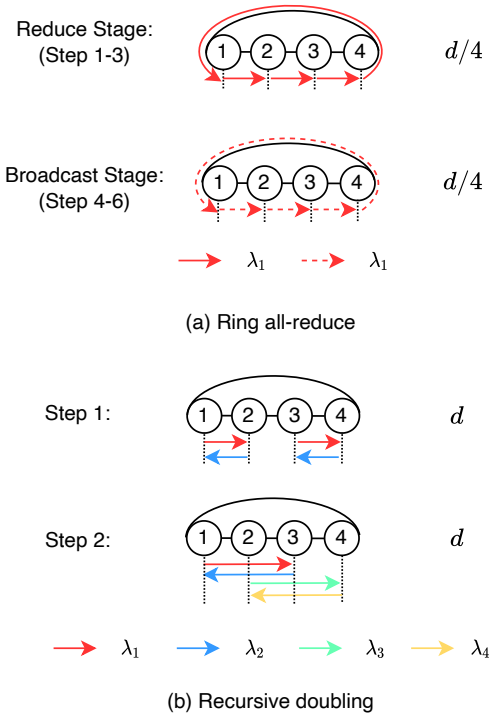

Fig. 3. Illustration of Ring and RD all-reduce operations in a 4-node optical interconnect system.

Table 1. Comparison of the number of communication steps for Ring and RD all-reduce algorithms in optical and electrical interconnect systems.

Algorithms	Network	Communication steps
Ring	Optical	$2(n-1)$
	Electrical	$2(n-1)$
RD	Optical	$\lceil log_2 n \rceil$, (n $\leq$ 2w) $2\lceil log_2 n \rceil - \lceil log_2 2w \rceil$, (n > 2w)
	Electrical	$\lceil log_2 n \rceil$

Communication time This subsubsection explores the communication time in data-parallel distributed DNN training, defined as the time required for all nodes to exchange data of size d with every other node. Given our previous assessment of the number of communication steps and the size of transferred data, we can now establish the communication time for both the Ring and RD all-reduce algorithms in the optical and electrical interconnect systems.

In an optical interconnect system, let a denote the reconfiguration time of the MRR and delay of O/E/O conversion, and β_O represent the bandwidth per wavelength. Based on Table 1, the communication time of Ring all-reduce in a system with n nodes for data-parallel distributed DNN training can be calculated by

$$T_{O-Ring} = 2(n-1)\left(\frac{d\eta}{n\beta_O} + a\right), \tag{6}$$

where η represents the storage size of one parameter.

Similarly, the communication time of RD all-reduce in an optical interconnect system with n nodes can be calculated as

$$T_{O-RD} = \theta\left(\frac{d\eta}{\beta_O} + a\right), \tag{7}$$

where θ is the number of communication steps when using the RD all-reduce algorithm.

We also formulate the communication time of the two all-reduce algorithms in an electrical interconnect system for data-parallel distributed DNN training. We use α to denote the link latency and β_E to denote the network bandwidth. According to Table 1, the communication cost of Ring all-reduce in a n nodes electrical interconnect system can be calculated as

$$T_{E-Ring} = 2(n-1)\left(\alpha + \frac{d\eta}{n\beta_E}\right). \tag{8}$$

Likewise, the communication cost of RD all-reduce in an electrical interconnect system with n nodes can be calculated as

$$T_{E-RD} = \lceil log_2 n \rceil \left(\alpha + \frac{d\eta}{\beta_E}\right). \tag{9}$$

5 Evaluation

This section details the methodologies used for simulating data-parallel distributed DNN training across both optical and electrical interconnect systems. In our approach, we developed a custom optical interconnect simulator in Python, aimed at illustrating the behavior of both Ring and RD all-reduce algorithms. The accuracy and reliability of this simulator were ensured through meticulous modeling and disciplined software development techniques. Given the lack of real-world optical interconnect hardware for direct validation, we implemented

additional safeguards. Specifically, we cross-verified the simulator's outputs with analytical and mathematical models to assure both accuracy and consistency. Though not tested against real-world configurations, our simulator is grounded in widely-accepted models of optical interconnects and machine learning frameworks. This simulated environment provides a thorough understanding of the relevant algorithms within the context of optical communications. For the electrical interconnect system, we replicated a conventional data center cluster environment. These typically employ a multi-layer Fat-tree topology and utilize electrical packet switches, as cited in [18]. This emulation was realized using SimGrid framework's version 3.3 discrete-event simulator, which is well-known for its effectiveness in modeling Fat-tree electrical interconnect systems [12]. It is worth noting that we used float32 as the data type for computations, and GPUs served as accelerators in both simulation scenarios.

To bolster the validity of our simulations, we used four widely-recognized DNN benchmarks: AlexNet, VGG16, ResNet50, and GoogLeNet. These models were strategically chosen for their compatibility with GPU memory constraints, a critical prerequisite for distributed data-parallel training since each GPU must house a full copy of the DNN model. To ensure a diverse set of test conditions, we executed these benchmarks using both the MNIST and ImageNet datasets, as cited in [11]. To precisely profile computation times, memory footprints, and data sizes for backpropagation, among other variables, we utilized the TensorFlow profiler [17]. Our test system for profiling was equipped with an Intel i7-6700K 4.0 GHz CPU, 64GB RAM, and a GeForce GTX TITAN XP GPU. This process yielded invaluable insights, such as optimal batch sizes and data transfer volumes for each of the four DNN models under varying numbers of GPUs. Notably, the profiler also verified the uniformity of gradient quantities—excluding the input and output layers—across multiple datasets within the same batch during distributed data-parallel training.

In terms of network-level simulations, we employed SimGrid to model the communications behavior of two key all-reduce algorithms in the electrical interconnect system designated for distributed DNN training. Table 2 outlines the simulated platform settings along with their associated parameters, all of which adhere to SimGrid's default configurations. Given that SimGrid does not natively support real DNN applications, we devised a specialized benchmark to simulate data-parallel distributed DNN training, as detailed in Algorithm 1. For simulations involving the optical interconnect system, we relied on parameters detailed in Table 2, which were sourced from references [8,22]. These parameters were implemented in our custom optical interconnect simulator. To ensure a rigorous performance evaluation, we benchmarked both the optical and electrical interconnect systems by comparing the efficiencies of the Ring and RD all-reduce algorithms. This comparison was conducted using both SimGrid and our in-house optical simulator. Estimates of communication time for these setups were derived through numerical computation, factoring in the number of parameters as calculated in both SimGrid and our custom simulator.

Algorithm 1: Pseudo-code of the benchmark

input : $N \leftarrow$ the number of nodes
$M \leftarrow$ the number of parameters
$I \leftarrow$ the number of iterations
output: Communication time of benchmark C

```
t1 ← MPI_Wtime()
Allocate buffers with M × sizeof(float) bytes
Initialize the value of send buf
for i ← 1 to I do
    for i ← 1 to M do
        MPI_Allreduce(send_buf, rec_buf, M, MPI_FLOAT, ...)
    end
end
C ← MPI_Wtime() - t1
```

5.1 Experiment Setup

Our experimental framework consists of two distinct sets of tests. The first set aims to investigate the impact of varying data transfer sizes on the Ring and RD all-reduce algorithms within both the electrical and optical interconnect systems. In the second set, we focus on the scalability of the two all-reduce algorithms in optical and electrical interconnect systems. We manipulate the number of nodes within our in-house simulator and SimGrid while maintaining the data transfer sizes of four representative DNN models. Additionally, we scrutinize the influence of varying wavelengths in the optical interconnect system, using a configuration of 1024 nodes. It is important to note that our evaluation of communication time is predicated upon one epoch of training time, given the repetitive nature of the training process.

Table 2. Parameters of simulated architecture.

Interconnect	Parameters setup
Eletrical network	Two-level fat-tree topology, Router bandwidth: 25 Gbps, Router delay: 25 μs, Packet size: 128 bytes, Shortest-path routing.
Optical network	Ring topology with two directions, 64 wavelengths per wavegiude, MRRs reconfiguration delay: 25 μs, O/E/O conversion delay: 1 cycle/flit, Packet size: 128 bytes, Flit size: 32 bytes, 40 Gbps/wavelength.

5.2 Simulation Results

Results of the first set We compared the performance of the Ring and RD all-reduce algorithms within the electrical and optical interconnect systems by deploying these algorithms in our in-house simulator and SimGrid. This configuration consisted of 1024 nodes and used transferred data sizes of 16MB, 64MB, 256MB, 1024MB, 4096MB, and 16384MB. For the optical interconnect system, we assumed 64 wavelengths.

Fig. 4 illustrates the performance of the Ring and RD all-reduce algorithms within optical and electrical interconnect systems consisting of 1024 nodes, with incrementally increasing data transfer sizes. A zoomed-in segment of the line charts that depict the performance of these algorithms on optical and electrical interconnect systems is presented in the inset of Fig. 4. From this inset, we can observe that the Ring all-reduce algorithm surpasses the RD all-reduce algorithm when data transfer exceeds 64MB within the optical interconnect system. This can be attributed to the trade-off between data transfer sizes and the number of communication steps in the Ring and RD all-reduce algorithms, with 64MB acting as a pivot point in the optical interconnect system. However, this pivot point differs within the electrical interconnect system. As shown in the inset of Fig. 4, the communication cost of the Ring all-reduce algorithm remains higher than the RD all-reduce algorithm until the transferred data exceeds 1024MB. This is due to the Ring all-reduce algorithm's larger number of communication steps and low bandwidth utilization in the electrical link when the data transfer is small, while the RD all-reduce algorithm's smaller number of communication steps and high bandwidth utilization of the electrical link.

Comparing the two interconnect systems, the optical interconnect system outperforms the electrical interconnect system for both the Ring and RD all-reduce algorithms. This superior performance of the optical interconnect system can be attributed to its higher bandwidth, lower latency, and smaller number of communication steps.

Results of the second set To evaluate the scalability of the two all-reduce algorithms in both interconnect systems, we scaled the nodes (128, 256, 512, and 1024) within the electrical and optical interconnect systems, employing transferred data sizes from four conventional DNN models. The batch sizes for AlexNet, VGG16, GoogLeNet, and ResNet50 were set at 512, 48, 64, and 1024, respectively, which are sizes that maximize GPU memory usage for all-reduce algorithms. All results in Fig. 5 are normalized by dividing by the result of O-RD in GoogLeNet.

Fig. 5 presents a comparison of the communication times of the Ring and RD all-reduce algorithms in the electrical and optical interconnect systems, with different DNN models across various scales. From Fig. 5 (a) to (d), it is apparent that: 1) the communication time of both all-reduce algorithms in the optical and electrical interconnect systems increases with the number of nodes, and 2) the communication time for both all-reduce algorithms is lower in the optical interconnect system compared to the electrical interconnect system. We observe

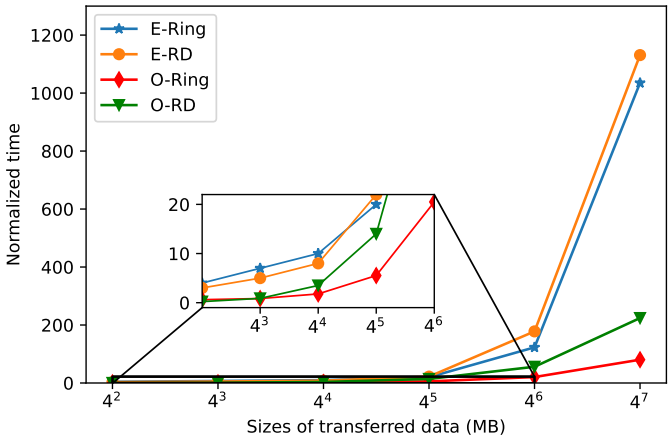

Fig. 4. Performance of the algorithms in the 1024 nodes optical and electrical interconnect system using different amount of transffred data.

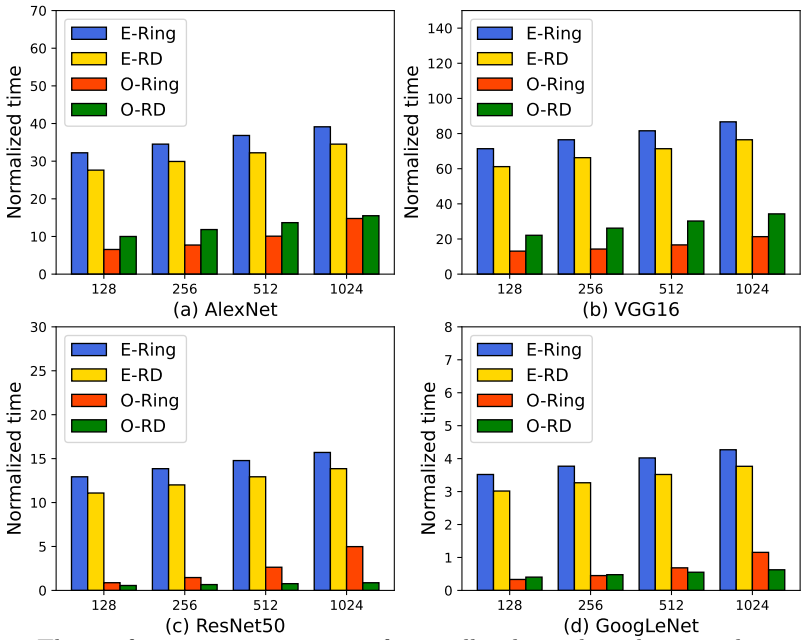

Fig. 5. The performance comparison of two all-reduce algorithms in electrical and optical interconnect systems.

the most significant change in communication time with the RD all-reduce algorithm in the optical interconnect system. The communication time for RD remains low and fairly constant with increasing nodes for DNNs with small data transfer sizes, such as ResNet50 and GoogLeNet. However, for DNNs with large data transfers, such as AlexNet and VGG16, the communication time for RD is higher than that of Ring all-reduce. This is because the communication time for RD grows proportionally with the data transfer size, as described in Eq. (7). Contrarily, in the electrical interconnect system, RD always outperforms Ring all-reduce due to RD's higher link utilization and relatively constant communication steps with the increase in nodes. Comparing the Ring all-reduce algorithm in optical and electrical interconnect systems, O-Ring significantly reduces communication time compared with E-Ring. This performance gain for the Ring all-reduce algorithm in the optical interconnect system is due to the system's higher bandwidth and lower latency. Overall, applying the RD and Ring all-reduce algorithms in the optical interconnect system reduces communication time on average by 74.74% and 75.35%, respectively, compared to the electrical interconnect system.

Furthermore, we evaluated the impact of wavelength on the performance of the two all-reduce algorithms in the 1024 node optical interconnect system, varying the number of wavelengths (8, 16, 32, 64, 128) with the data sizes from the four traditional DNN models. Fig. 6 displays the performances of the two all-reduce algorithms using different numbers of wavelengths for the four DNNs. As the number of wavelengths increases, the communication time of RD decreases while Ring all-reduce remains constant. This suggests that the communication cost of RD is related to the number of wavelengths, while Ring is not.

From the above simulation results, we can draw the following conclusions: 1) In the 1024-node optical and electrical interconnect systems, Ring all-reduce outperforms RD when the data transfer size exceeds 64MB and 1024MB, respectively. 2) The optical interconnect system exhibits superior scalability over the electrical interconnect system. Both Ring and RD all-reduce in the optical interconnect system reduce average communication time by 74.74% and 75.35%, respectively, compared to the electrical interconnect system across all four DNNs. 3) In the optical interconnect system, the communication time of Ring is unaffected by the number of wavelengths, while the communication time of RD decreases as the number of wavelengths increases.

5.3 Discussion

In light of our findings, we believe that a hybrid approach to implementing all-reduce algorithms could significantly improve distributed DNN training performance in both electrical and optical interconnect systems. This approach would allow us to tailor the algorithm's behavior to the specific characteristics of the system and the size of the data being transferred.

In the context of optical interconnect systems, we observed that neither the Ring nor the RD all-reduce algorithms fully exploit the available wavelengths at each step. For example, Ring all-reduce uses only one wavelength per step, and

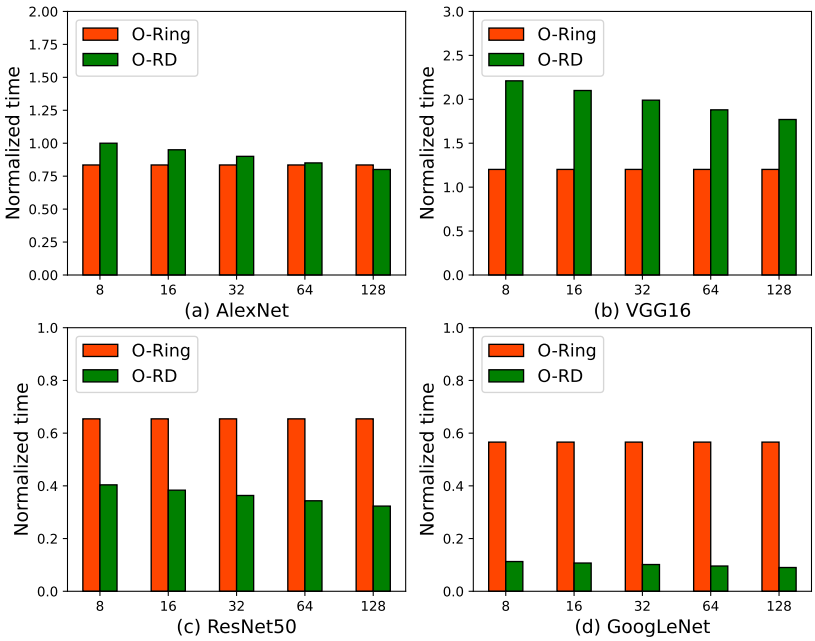

Fig. 6. The wavelength impact on Ring and RD all-reduce in the 1024 nodes optical interconenct systems.

although RD can utilize multiple wavelengths, its usage is unbalanced and does not reuse wavelengths to further minimize the number of communication steps.

Furthermore, we identified a trade-off between the number of communication steps and the size of the data transferred in all-reduce algorithms. In the Ring all-reduce algorithm, the amount of data transferred per step is small ($\frac{d}{n}$), but the number of steps is large ($2(n-1)$), which linearly increases as the number of nodes increases. On the other hand, in RD, the amount of data transferred per step is larger (d), but the number of steps is smaller ($2\lceil log_2 n \rceil - \lceil log_2 2w \rceil$, where $n > 2w$).

To further enhance communication efficiency, the all-reduce algorithm in the optical interconnect system could be customized. The goal would be to fully utilize the wavelengths available at each communication step while keeping communication traffic as low as possible. This could significantly reduce the communication time in distributed DNN training, contributing to improved overall system performance.

6 Conclusion

In this paper, we undertook a comparative analysis of Ring and RD all-reduce algorithms within optical and electrical interconnect systems, with the goal of better understanding their performance characteristics. After formulating the

communication cost of the two all-reduce algorithms, we carried out extensive simulations to further compare their performance under various system settings. Our findings can be summarized as follows: 1) The Ring all-reduce algorithm demonstrated superior performance over RD when transferring data exceeding 64MB and 1024MB in optical and electrical interconnect systems with 1024 nodes, respectively. 2) The optical interconnect system proved to have better scalability than the electrical interconnect system. Furthermore, both Ring and RD all-reduce operations within the optical interconnect system experienced an average communication time reduction of 74.74% and 75.35% compared to the electrical interconnect system across four distinct DNNs. 3) Within the optical interconnect system, the communication time of the Ring algorithm did not show a dependency on the number of wavelengths. However, the communication time of the RD algorithm could be reduced with an increase in the number of wavelengths. Our study offers a foundation for future research to focus on optimizing all-reduce algorithms for distributed DNN training to further reduce communication time in optical interconnect systems. Moreover, exploring the performance differences of these algorithms in different topologies could offer additional insights to enhance system performance in distributed DL applications.

Acknowledgements We thank the reviewers for taking the time and effort necessary to review the manuscript. Besides, we acknowledge using New Zealand eScience Infrastructure (NeSI) high-performance computing facilities as part of this research (Project code: uoo03633).

References

1. Khan, A. R., Kashif, M., Jhaveri, R. H., Raut, R., Saba, T., Bahaj, S. A.: Deep learning for intrusion detection and security of Internet of things (IoT): current analysis, challenges, and possible solutions. Security and Communication Networks 2022 (2022).
2. Luo, L., West, P., Nelson, J., Krishnamurthy, A., Ceze, L.: Plink: Discovering and exploiting locality for accelerated distributed training on the public cloud. Proceedings of Machine Learning and Systems, 2, pp. 82–97 (2020).
3. Wang, G., Venkataraman, S., Phanishayee, A., Devanur, N., Thelin, J., Stoica, I.: Blink: Fast and generic collectives for distributed ml. Proceedings of Machine Learning and Systems, 2, pp. 172-186 (2020)
4. Yuichiro, U., Yokota, R.: Exhaustive study of hierarchical allreduce patterns for large messages between GPUs. In 2019 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), pp. 430-439 (2019)
5. Jiang, Y., Gu, H., Lu, Y., Yu, X.: 2D-HRA: Two-dimensional hierarchical ring-based all-reduce algorithm in large-scale distributed machine learning. IEEE Access, 8, 183, pp. 488–183, 494 (2020)
6. Cho, M., Finkler, U., Serrano, M., Kung, D., Hunter, H. . Blueconnect: Decomposing all-reduce for deep learning on heterogeneous network hierarchy. IBM Journal of Research and Development, 63(6), pp. 1–1 (2019)

7. Nguyen, T. T., Takano, R.: On the feasibility of hybrid electrical/optical switch architecture for large-scale training of distributed deep learning. In 2019 IEEE/ACM Workshop on Photonics-Optics Technology Oriented Networking, Information and Computing Systems (PHOTONICS), pp. 7–14 (2019)
8. Khani, M., Ghobadi, M., Alizadeh, M., Zhu, Z., Glick, M., Bergman, K., Vahdat, A., Klenk, B., Ebrahimi, E.: SIP-ML: High-bandwidth optical network interconnects for machine learning training. In Proceedings of the 2021 ACM SIGCOMM 2021 Conference, pp. 657–675 (2021).
9. Gu, R., Qiao, Y., Ji, Y.: Optical or electrical interconnects: Quantitative comparison from parallel computing performance view. In IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference pp. 1-5 (2008)
10. Shin, J., Seo, C. S., Chellappa, A., Brooke, M. A., Chatterjee, N. M., Jokerst.: Comparison of electrical and optical interconnect. In IEEE Electronic Components and Technology Conference pp. 1067-1072 (1999)
11. Wei, J., Ibrahim, Y., Qian, S., Wang, H., Liu, G., Yu, Q., et al.: Analyzing the impact of soft errors in VGG networks implemented on GPUs. Microelectronics Reliability 110, 113648 (2020)
12. Casanova, H., Legrand, A., Quinson, M.: SimGrid: A generic framework for large-scale distributed experiments. In Tenth IEEE International Conference on Computer Modeling and Simulation (UKSim2008) pp. 126–131 (2008)
13. Alotaibi, S., Dlaim, K., Yadav, A. N., Aledaily, L. M., Alkwai, et al.: Deep Neural Network-Based Intrusion Detection System through PCA. Mathematical Problems in Engineering (2022)
14. Huang, J., Majumder, P., Kim, S., Muzahid, A., Yum, K. H., Kim, E. J.: Communication Algorithm-Architecture Co-Design for Distributed Deep Learning. In 2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA). IEEE, pp. 181–194 (2021)
15. Ghobadi, M. . Emerging Optical Interconnects for AI Systems. In IEEE 2022 Optical Fiber Communications Conference and Exhibition (OFC), pp. 1-3 (2022)
16. Dai, F., Chen, Y., Huang, Z., Zhang, H., Zhang, F.: Efficient All-Reduce for Distributed DNN Training in Optical Interconnect Systems. In Proceedings of the 28th ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming, pp. 422-424 (2023)
17. TensorFlow, (n.d.). Optimize TensorFlow performance using the Profiler. Retrieved: September 2, 2023, from https://www.tensorflow.org/guide/profiler
18. Wang, W., Khazraee, M., Zhong, Z., Ghobadi, M., Jia, Z., Mudigere, D., et al.: TopoOpt: Co-optimizing Network Topology and Parallelization Strategy for Distributed Training Jobs. In 20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23), pp. 739-767 (2023).
19. Zhang, H., Zheng, Z., Xu, S., Dai, W., Ho, Q., Liang, X., et al.: Poseidon: An efficient communication architecture for distributed deep learning on GPU clusters. In 2017 USENIX Annual Technical Conference (USENIX ATC 17) pp. 181-193 (2017).
20. Dai, F., Chen, Y., Huang, Z., Zhang, H., Zhang, H., Xia, C.: Comparing the performance of multi-layer perceptron training on electrical and optical network-on-chips. The Journal of Supercomputing, 79(10), pp. 10725-10746 (2023).
21. Ottino, A., Benjamin, J., Zervas, G.: RAMP: A flat nanosecond optical network and MPI operations for distributed deep learning systems. Optical Switching and Networking, 100761. (2023).
22. Dai, F., Chen, Y., Zhang, H., Huang, Z.: Accelerating fully connected neural network on optical network-on-chip (onoc). arXiv preprint arXiv:2109.14878 (2021).

23. Xia, C., Chen, Y., Zhang, H., Zhang, H., Dai, F., Wu, J.: Efficient neural network accelerators with optical computing and communication. Computer Science and Information Systems 20(1), pp. 513-535 (2023).