Article

A Novel Decomposed Optical Architecture for Satellite Terrestrial Network Edge Computing

Xiaotao Guo [1] , **Ying Zhang** [2,*], **Yu Jiang** [1] , **Shenggang Wu** [1] **and Hengnian Li** [1]

[1] State Key Laboratory of Astronautic Dynamics, Xi'an Satellite Control Center, Xi'an 710043, China; guoxtchina@163.com (X.G.); jiangyu_xian_china@163.com (Y.J.); s_wu_xscc@163.com (S.W.); henry_xscc@mail.xjtu.edu.cn (H.L.)

[2] Information and Navigation School, Air Force Engineering University, Xi'an 710082, China

[*] Correspondence: y_zhang0505@163.com

Abstract: Aiming at providing a high-performance terrestrial network for edge computing in satellite networks, we experimentally demonstrate a high bandwidth and low latency decomposed optical computing architecture based on distributed Nanoseconds Optical Switches (NOS). Experimental validation of the decomposed computing network prototype employs a four-port NOS to interconnect four processor/memory cubes. The SOA-based optical gates provide an ON/OFF ratio greater than 60 dB, enabling none-error transmission at a Bit Error Rate (BER) of 1×10^{-9}. An end-to-end access latency of 122.3 ns and zero packet loss are obtained in the experimental assessment. Scalability and physical performance considering signal impairments when increasing the NOS port count are also investigated. An output OSNR of up to 30.5 dB and an none-error transmission with 1.5 dB penalty is obtained when scaling the NOS port count to 64. Moreover, exploiting the experimentally measured parameters, the network performance of NOS-based decomposed computing architecture is numerically assessed under larger network scales. The results indicate that, under a 4096-cube network scale, the NOS-based decomposed computing architecture achieves 148.5 ns end-to-end latency inside the same rack and zero packet loss at a link bandwidth of 40 Gb/s.

Keywords: satellite-terrestrial network; hardware decomposition; nanoseconds optical switches

MSC: 78-05; 94-10

1. Introduction

In order to satisfy various requirements of emerging big data applications (like Internet of Things (IoT), automatic driving, and high-quality video/image processing), a novel integrated Satellite-Terrestrial Network (STN) has been proposed combing the space-based satellite communication system and terrestrial computing network [1–4]. The STN achieves stable global coverage to the user-terminal utilizing satellite networks, while relieving the limited processing performance of the satellite system by offloading the computing procedure to terrestrial computing infrastructures [5,6]. The satellites can be applied as front-end nodes and process delay sensitive tasks with less requirements of computing performance. Meanwhile, the terrestrial network coordinates with satellites as the back-end to process computing intensive tasks. This architecture requires a powerful and efficient terrestrial computing network. However, because of the integrated on-board hardware in current server-centric terrestrial computing architecture, a large amount of hardware resources (processor, memory, and storage) is frequently underutilized [7]. Moreover, the application performance can be degraded under server-centric computing architectures when limited on-board hardware resources cannot meet the application requirements. In addition, the failure or upgrade of just one kind of resource (only processor or memory) in a server causes a significant impact to the whole server.

To relieve the above issues of server-centric architectures, a novel computing architecture consisting of decomposed hardware cubes (hence, naming decomposed architecture) is developed to provide flexible resource provision and boosting performance for various processor (memory) intensive applications [8–12]. In order to implement the decomposed architecture, there are several issues to be addressed, in which the disaggregation of the processor and memory is a key challenge. Due to the hardware decomposition, the on-board high-speed point-to-point bus between the processor and memory is replaced by the network interconnection. This difference requires a network connection of high bandwidth and low latency to sustain the traffic among decomposed hardware cubes. The network switching may also lead to signal deterioration (including channel crosstalk, OSNR degradation, signal distortion, and power budget), hence, degrading the computing performance. Therefore, a stable, error-free, and low packet loss network is also significantly important for decomposed architectures.

In this paper, we design and implement an optical terrestrial computing network with decomposed hardware for STN edge computing, while hardware cubes are interconnected by an NOS-based flat network to minimize the communication delay. An independent optical packet control plane is designed to process the optical packet switching and solve the packet contention. The decomposed computing prototype is experimentally implemented with a 4 × 4 NOS and four processor/memory cubes. The hardware resource cubes (processor and memory) are implemented based on Field-Programmable Gate Array (FPGA) platforms. The physical and network performance of the optical computing prototype with decomposed hardware are experimentally assessed, while the scalability of the decomposed computing architecture (up to 4096 hardware cubes) is also numerically investigated based on experimentally measured parameters.

The article is organized as follows. Section 2 provides an overview of the existing works and describes the contributions of this work. In Section 3, the NOS-based decomposed optical computing architecture for STN and its operation are presented. In Section 4, the experimental demonstration of the NOS-based decomposed computing prototype is discussed. Section 5 shows the scalability of the NOS-based computing architecture with decomposed hardware in terms of physical performance and network performance. In addition, our proposed optical computing network with decomposed hardware is compared with current works in this section. The main results of this work are concluded in Section 6.

2. Literature Review and Contribution

2.1. Related Works

STN has been investigated as a promising solution to provide various cloud services and applications in next generation network architectures. Xie et al. [2] deployed the computing resource in multi-layer heterogeneous edge computing clusters for STN, considering the Quality of Service (QoS) requirements, computation offloading, and task scheduling. Wang et al. [3] designed a space edge computing node for STN to provide services coordinated with the terrestrial computing network, taking less computing time and consuming less energy than traditional satellite constellation. Tang et al. [5] presented a three-tier computation architecture combining hybrid terrestrial cloud computing architecture and an edge computing Low Earth Orbit (LEO) satellite network to minimize the sum power consumption of the user-terminal. Pfandzelter et al. [13] proposed a resource place method to select a subset of satellites as computing nodes with QoS constraints. A joint offloading decision and resource allocation scheme is designed in [14] considering satellites as the computing node to assist the terrestrial computing network in task processing, which minimizes the completion delay of computing tasks.

In order to implement the STN in the above studies and provide a minimal delay for the user-terminal with QoS requirements, a powerful terrestrial computing network is necessary. The decomposed computing architecture has been considered as an emerging paradigm with flexible resource provision and faster task runtime. Several experimental studies have been conducted to investigate the feasibility of decomposed architectures.

According to the analysis in [15], the low packet loss and long-term network stability are critical, but the extra network latency impacts the performance of decomposed architectures the most. Some decomposed architectures were proposed based on current servers and the hierarchical electrical network to implement decoupled memory blades [16,17]. A decomposed in-memory store framework was proposed in [18] for big data applications, utilizing theThymesisFlow memory system. A Software-Defined Networking (SDN)-based orchestration plane is designed and experimentally implemented for a decomposed computing architecture in [19] with hybrid optical switches. Decomposed architectures based on optical switches were proposed and experimentally investigated in [20,21], exploiting high aggregation bandwidths and transparent switching; however, the high switching delay (milliseconds) of the optical switch technology cannot provide nanoseconds scale communication between processor and memory cubes. An optically connected memory architecture was demonstrated in [22,23] based on micro-ring resonator switches and the Aurora 64B/66B protocol, which achieved an optical switching time of hundreds of microseconds. To provide fast optical switching, a flat optical network was developed based on tunable lasers and passive gratings [24], which may lead to unstable operation during high-speed processing. A decomposed optical computing architecture was proposed based on Nanoseconds Optical Switches (NOS) [25]. The NOS is based on a broadcast and select switch architecture employing Semiconductor Optical Amplifiers (SOA)-based optical gates. Exploiting the nanoseconds switching of SOA-based gates and parallel optical flow control, the proposed architecture can potentially provide an optical network interconnection of high bandwidth and low latency for decomposed architectures.

2.2. Motivation and Contribution

Despite the fact that the decomposed computing architecture can provide a more flexible and efficient terrestrial network with faster application runtime for STN, the interconnection and communication among hardware cubes are still challenges, which requires low access latency and high aggregated bandwidth. Existing electrical network-based solutions is based on commercially available hardware, which needs to be compatible with the current on-board bus protocol. This inherent high latency due to the on-board peripheral I/O connection and hierarchical server-to-server interconnect network leads to degraded application performance, while repeatedly photovoltaic conversions deteriorate the energy efficiency of decomposed computing architectures. The optical switches-based network can avoid photovoltaic conversions and achieve a high aggregated bandwidth of up to hundreds of gigabits. However, the long reconfiguration time (up to milliseconds) limits the flexibility of the decomposed computing architectures. A stable and fast switching interconnect network is significant for the decomposed terrestrial computing architecture of STN. Our previous work in [25] only numerically assessed the application performance of NOS-based decomposed computing architecture, and the impact of optical network switching based on NOS as well as network performance must be experimentally investigated to validate the feasibility and scalability of the proposed decomposed architecture.

The contributions of our work are summarized as follows:

(1) A four-node decomposed computing prototype is experimentally implemented in this work, consisting of two processor nodes, two memory nodes, and one four-port NOS. The hardware cubes are implemented utilizing FPGA chip, and two independent interconnect channels are designed between hardware cubes and NOS for sending optical payload and signal tags respectively.

(2) The physical and network performance of the decomposed computing prototype are investigated in experimental assessments. In the physical assessment, the SOA-based optical gates achieve ON/OFF switch ratios larger than 60 dB and ensure low interchannel optical signal interference. The implemented prototype provides a none-error transmission among decomposed hardware cubes with 0.5 dB power compensation at a BER of 1×10^{-9}. Meanwhile, in the network performance evaluation, the prototype performs an end-to-end latency of 122.3 ns and zero packet loss after link establishing.

(3) As the NOS port count directly impacts the scalability and feasibility of decomposed architectures, we further investigate the physical performance in terms of output OSNR and required power penalty as a function of the NOS port count. The proposed decomposed architecture provides an output OSNR of up to 30.5 dB under the NOS port count of 64. Scaling the NOS port count to 64, an error-free operation with a power penalty of 1.5 dB is achieved.

(4) The scalability of the NOS-based computing network with decomposed hardware is also evaluated in this work. Based on the experimentally measured parameters, the network performance of the NOS-based decomposed architecture is also numerically assessed under different network scales and link bandwidths. The results show that with a scale of 4096 hardware cubes and a memory cube access rate of 0.9, an end-to-end latency of 148.5 ns inside a rack and an end-to-end latency of 265.6 ns across racks are obtained under a link bandwidth of 40 Gb/s.

3. Decomposed Optical Network for STN Edge Computing

The satellite-terrestrial edge computing network is shown in Figure 1a with a decomposed optical network. The satellite system receives the requests from user terminals and transfers corresponding instructions to terrestrial decomposed computing architecture via downlink. The decomposed optical computing architecture consists of diverse resource pools such as processor, memory (Mem), and GPU, which can be flexibly configured based on application requirements. The hardware-to-NOS interconnection is divided into two parallel parts: optical packet sending plane and control plane. In the packet sending plane, the NOS switch port is linked to hardware cubes for sending optical payloads; while, in the packet control plane, the packet tag from the hardware cube is processed by the switch controller to forward/block optical payloads and reduced optical packet contention. It is significant for the scalability of the decomposed network to implement an NOS with large port count.
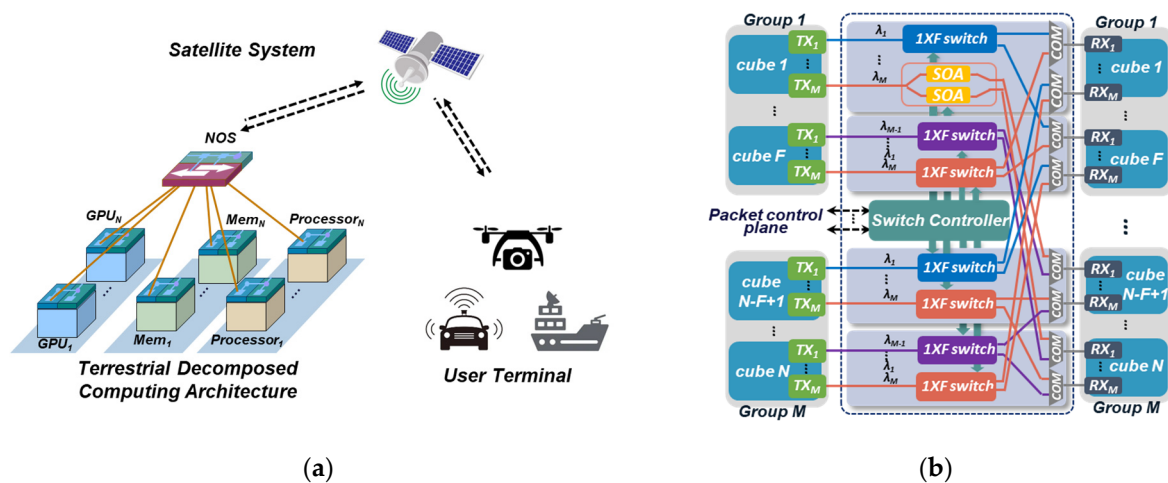

(a)   **(b)**

Figure 1. (**a**) Satellite terrestrial network with decomposed computing architecture; (**b**) Schematic of nanoseconds optical switch.

The schematic of NOS is depicted in Figure 1b. There are *N* hardware cubes in total, split into *M* groups and *F* hardware cubes in each group. Leveraging distributed processing modules, the NOS can process multiple optical packets from different hardware cubes in parallel. When the FPGA-implemented switch controller analyzes the packet tag from the control-path channel, the packet payload from the data-path channel is broadcasted to SOA-based gates of the $1 \times F$ switch using a splitter. The hardware cube grouping can scale the port count of the NOS utilizing multiple $1 \times F$ switch with smaller port counts. SOA

gates can provide nanoseconds switching time and compensate the splitting power loss. The number of cells N_{cell} that a packet payload occupied in TX is calculated as follows:

$$N_{cell} = \lceil L_p / L_b \rceil \tag{1}$$

where L_p and L_b are the length of packet and each buffer unit, respectively. After the $1 \times F$ switch, one $F \times 1$ Arrayed Waveguide Gratings (AWG) gathers F different optical wavelengths from the same group and sends packets to the corresponding receiver. According to the packet destination information in the packet tag, the controller switches the corresponding SOA gate on and rest optical gates off, then the packet payload is forwarded to the target hardware cube. Taken the group that wavelength λ_i of the hardware cube n ($1 \leq n \leq F$, $1 \leq I \leq M$) destine to as d. When M is equal to or larger than F, a feasible wavelength mapping rule is shown in the following formula:

$$d = (i + 1 - n) \bmod M \tag{2}$$

Due to the processing delay of the switch controller in the control-path channel, hardware cubes send out optical packets with the respective delay. To guarantee the alignment of the tag and packet, there is a periodic calibration in the control-path channel. The switch controller is also responsible for solving potential contentions among optical packets. If packet contention occurs, the optical payload with higher priority is transmitted at first, while the rest of the packet payloads sent in the same time slot are blocked by switching corresponding SOA gates off. The switch controller then sends the acknowledgment signals to the corresponding hardware cubes for successfully forwarding the packet (ACK) or re-transmitting the packet (NACK) [26]. Benefiting from the structure of the distributed modules and nanoseconds switching time of SOA gates, the NOS can provide a scalable network with the minimal network latency and high aggregation bandwidth for the decomposed hardware cubes interconnection. More details about the NOS can be found in [27].

Each hardware cube has three kinds of components: on-board resource, resource management module, and network interconnect module. The functional diagram of the processor cube is illustrated in Figure 2a. The on-board resource of the processor cubes is the CPU, and a small memory chunk (local memory) is still placed inside the processor cube to run the operating system and cache core data. Both the CPU and local memory are interconnected to the Memory Management Unit (MMU), which is the resource management module of the processor cube. The MMU is responsible to the translation of virtual data addresses to physical data addresses. The network interconnect module consists of a flow controller, Network Interface (NI), and transmitters (TX and RX in Figure 2a,b). The flow controller is for sending optical packets to corresponding TX and processing packets from RX, while NI can package the instructions to network packets. All the TRXs are divided into two parts: some in the packet plane for sending payloads and others in the control plane for sending optical tags. The functional diagram of the memory cube is shown in Figure 2b. The on-board resource of the memory cubes is based on the Double Data Rate 4 (DDR4) memory or Hybrid Memory Cube (HMC), connected to the memory controller (resource management module of memory cube). The memory cubes have the same network interconnect module as processor cubes, processing packets of reading/writing data from processor cubes and packets from other memory cubes for the direct memory access. Same as current server-centric architectures, the CPU first accesses the fast cache when processing application data [28]. Once data are missing in the fast cache, the instruction and logical data address are sent to the MMU. If the physical address of the target data locates in the local memory, CPU accesses the target data within the processor cube. Otherwise, the instruction and physical data address are forwarded to the flow controller. Based on the hardware cube address containing the target data address, the instruction is packaged in NI, and then sent to the corresponding TX and forwarded to destination memory cube.
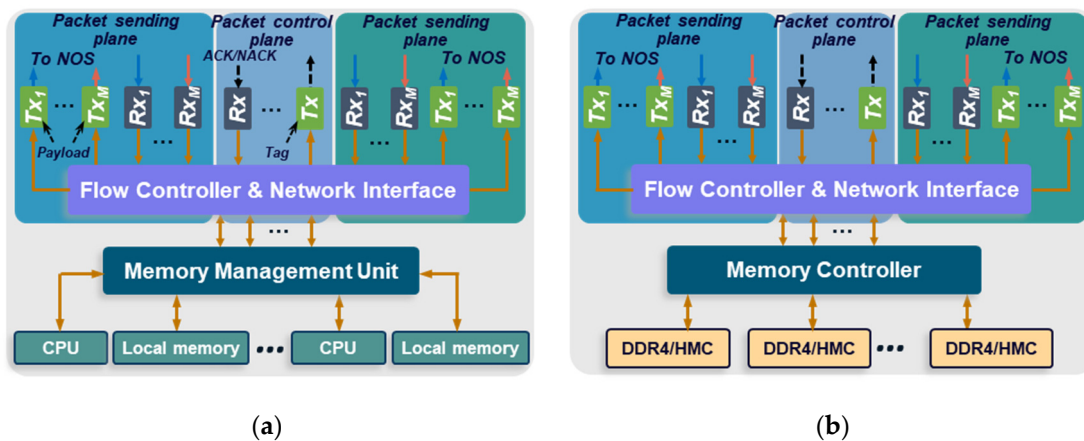
(**a**)   (**b**)

Figure 2. Functional diagram of (**a**) processor cube and (**b**) memory cube.

4. Experimental Demonstration of Decomposed Optical Architecture

4.1. Experimental Setup

The implemented experimental computing prototype with decomposed hardware is depicted in Figure 3. There are four FPGA-based hardware cubes (two processor cubes and two memory cubes) with one four-port-count NOS in the prototype. Four hardware cubes are designed via Vivado utilizing Xilinx Vertex VU095 [29], and the controller for the four-port-count optical switch is implemented with Xilinx Vertex-7 VC709 [30]. In this work, we use DDR4 memory as the on-board resource of memory cubes. All the four hardware cubes of NOS interconnections are based on two parallel commercial 10 Gb/s SFP+ transceivers (TRX). One TRX is to connect hardware cube and the NOS input/output port for the optical packet sending plane, while another TRX is for the optical packet control plane to process the packet tag. All the components in the hardware cube (including MMU, flow controller, and NI) are designed by FPGA circuit with a 322.3 MHz reference clock for data processing. In the processor cube, all the transferred data are handled via 32 bit-width format, and the payload length of each packet is set to 512 bits, as this value is the classical size of a cache line in existing computing structures. Before sending the optical payload, the packet has an eight-bit size preamble code to set up the link between hardware cube and NOS. Meanwhile, there is also an eight-bit size check code after the optical payload. Between two continuous optical packets, it has a slot to identify the start/end of packets, including 16 ns slack and 6 ns rise/fall times. The transmitter has an output of 1.2 dBm by utilizing the SFP+ transceiver, while a 2 m-length fiber is applied to connect the hardware cubes and NOS. Each SOA-based optical gate is interconnected to the input/output port of the NOS with a 1 m-length fiber. In the packet sending plane of the four-port-count NOS, there are four 1-to-4 optical signal splitters, sixteen SOA-based optical gates, and four 4-to-1 optical signal couplers in total. The switch controller distributes its reference clock to all the hardware cubes for time synchronization, and the processing of the optical packet sending plane and control plane is synchronized by utilizing the same clock and periodic calibration. At the same time, blank packets are sent to keep the link alive during periods of no data packets.
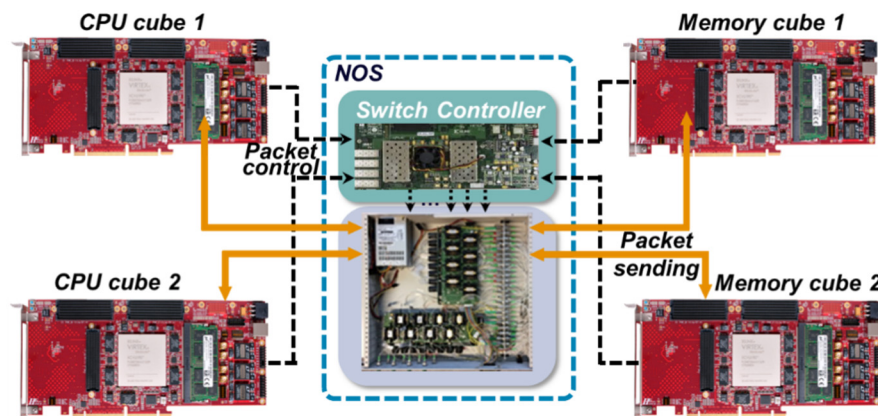
Figure 3. Experimental setup consisting of four hardware cubes and one NOS.

4.2. Physical Performance of Decomposed Prototype

Due to the broadcast and selective structure of NOS, only one of *N* optical gates is set to ON status during sending optical payloads to the corresponding destination hardware cube. This procedure requires the rest of the optical SOA gates to keep any optical signal from passing NOS. Otherwise, these leaked optical signals may lead to the signal interference after optical signal coupling at the output port of NOS. Therefore, in order to evaluate the implemented prototype and quantify the performance degradation due to optical signal interference, the optical power difference of SOA-based optical gates switching between ON and OFF status (ON/OFF ratio) is assessed in this part. It is depicted in Figure 4a that the SOA-based optical gate obtains a larger power difference when switching between ON and OFF status at higher SOA operation currents (better performance). When increasing the SOA operation current higher than 30 mA, the power difference of SOA-based optical gates is larger than 60 dB. The power spectra of optical signal passing through the SOA-based optical gates are shown in Figure 4b with ON and OFF status, respectively, and SOA operation current is configured as 60 mA in the assessment. With an ON/OFF ratio of higher than 60 dB, the SOA gates can block almost all the optical signals at OFF status, avoiding extra noise at output channel. Meanwhile, the target optical signal is amplified though ON SOA gate (output channel), guaranteeing a minimal signal interference across channels. Benefiting from the operation features of SOA, we can set the ON/OFF status of SOA gates within nanoseconds and achieve a rapid packet switching with good signal quality.
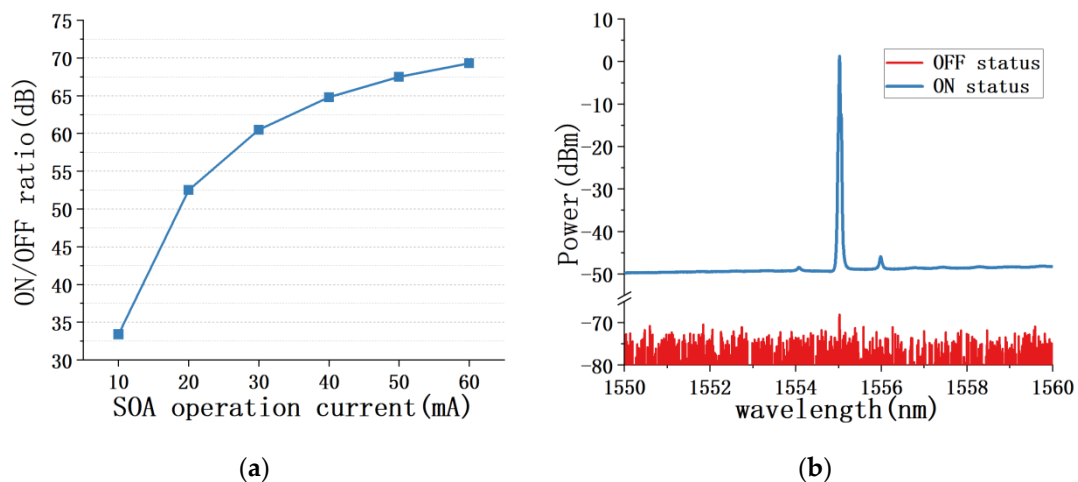

(**a**)

(**b**)

Figure 4. (**a**) Optical signal power difference with ON and OFF SOA status; (**b**) Power spectra of optical signal passing through SOA based optical gates.

To quantify the physical performance and validate the feasibility of the decomposed computing architecture with decomposed hardware, the BER of the implemented computing prototype is assessed with a SOA operation current of 60 mA, as shown in Figure 5. The Xilinx IBERT module [31] is deployed in all the four hardware cubes for calculating the BER of the computing prototype. BER is recorded when processor cubes send instructions to memory cubes. The commercially available plug-in SFP+ transceivers are applied in the assessment. The end-to-end BER without NOS (B-to-B) is measured as a criterion in the evaluation. It is shown in Figure 5 that the implemented computing prototype can perform the none-error packet transmission with 0.5 dB power compensation under a BER of 1×10^{-9} and operation current of 60 mA. This demonstrates that the implemented computing prototype can provide a feasible physical interconnection among decomposed hardware cubes
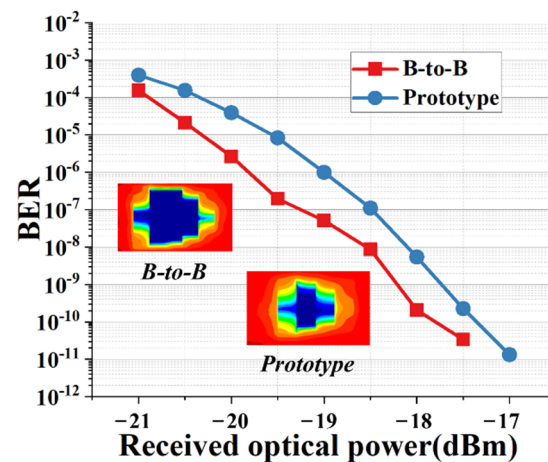

Figure 5. BER and eye diagram of decomposed computing prototype.

4.3. Network Performance of Decomposed Computing Prototype

The network performance of the decomposed computing prototype is assessed in terms of end-to-end access latency and packet loss. The network latency of the decomposed prototype includes four components: the processing delay in the flow controller and network interface of the source and destination hardware cubes, packet switching in the NOS, and fiber transmission delay. The end-to-end access latency of the implemented computing prototype with its breakdown is illustrated in Figure 6a, in which the latency is recorded when the processor cubes send instructions to the memory cubes. The customized transmission protocol is designed for the computing prototype with decomposed hardware to minimize the protocol processing delay. It is shown that, based on customized transmission protocol, the processor cube contributes 20.63 ns to handle data access instructions before sending optical packets out, while the memory cube takes 28.27 ns for receiving instructions and finding the target data. The NOS switching delay is higher compared to other components because it has more processing procedure including control signal processing, switch driver delay, and acknowledgement signal generation. Utilizing fast packet switching by NOS and customized transmission protocol for decomposed hardware, the implemented computing prototype performs minimal end-to-end access latency of 122.3 ns. By applying the Application Specific Integrated Circuit (ASIC) to implement the function of the decomposed hardware instead of the FPGA chip, the end-to-end access latency can be reduced still further.
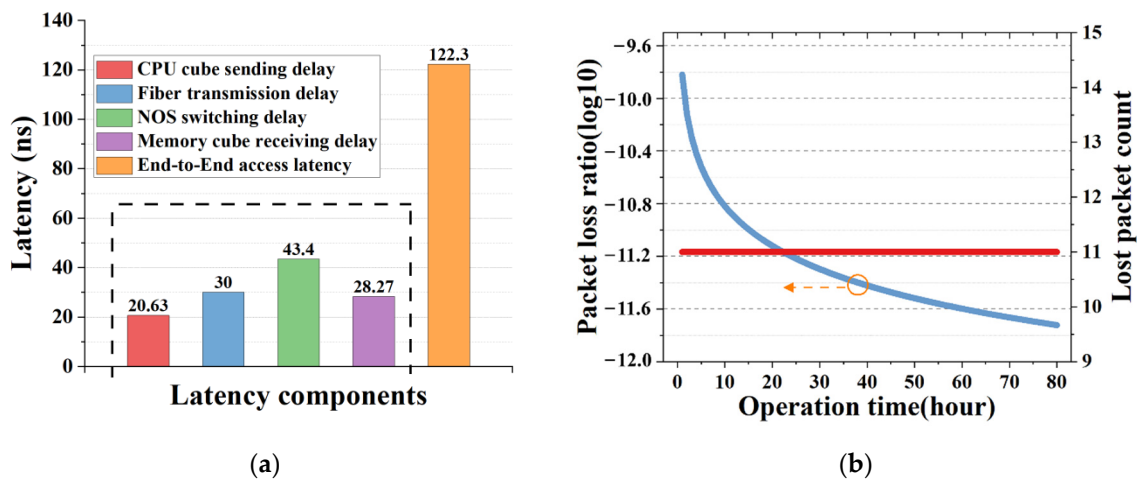
(**a**)　　　　　　　　　　　　(**b**)

Figure 6. (**a**) Network latency components when accessing memory cube; (**b**) Cumulative network packet loss over 80 h.

To evaluate the network transmission stability of the computing prototype with decomposed hardware, the packet loss is cumulatively measured during 80-h operation time. In the assessment, the processor cube sent 5.8×10^{-11} instructions to memory cube in total. As shown in Figure 6b, the cumulative packet loss of the implemented computing prototype with decomposed hardware is less than 1.9×10^{-12} in the end. Only 11 packets are lost during initialization stage before setting up the alive link. This is because the Clock and Data Recovery (CDR) procedure is required during initialization stage at receiver part. After finishing the initial CDR procedure, the NOS based network can provide a stable and no packet loss interconnection for the decomposed hardware cubes. This is because that the link keeps active and the clock is locked by inserting blank packets, avoiding resetting up the link.

5. Scalability and Discussion

The decomposed architecture prototype has been experimentally validated in Section 4, consisting of four hardware cubes. To better investigate the scalability of decomposed optical computing architecture, the physical and network performance of decomposed architecture are evaluated in this section under different network scales. For the physical performance evaluation, the output OSNR and power penalty for error-free operation are experimentally assessed as a function of the NOS port count. Meanwhile, exploiting the experimentally measured parameters, the network latency of decomposed computing architecture is numerically investigated under different network scales and bandwidths.

5.1. Physical Performance under Different Port Counts

In the optical computing architecture with decomposed hardware, all the on-board interconnect buses are removed, and NOS based flat optical network is applied to interconnect hardware cube. Thus, the port count of NOS is significantly important for the network scale of optical computing architecture with decomposed hardware. Due to the broadcast & select structure, the input power of OSA based optical gates is dramatically reduced when port count of NOS increases. Therefore, the output Optical Signal Noise Ratio (OSNR) of NOS is evaluated with a range of NOS port counts (four-port to 64-port). The output OSNR of decomposed hardware cubes is 62.3 dB in the experiment. With four different SOA operation currents (60 mA, 90 mA, 120 mA, and 150 mA in the experiment), the received OSNR at the decomposed hardware side is measured by increasing the NOS port count from 4 to 64. It is shown in Figure 7a that, with SOA-based optical gates operated at 150 mA, the decomposed hardware cube receives optical signals of 43.3 dB OSNR utilizing a four-port NOS, and obtains optical signals of 30.5 dB OSNR utilizing a 64-port NOS. In addition, it is shown that, when deploying the NOS with small port counts in the

experiment, the current values applied for SOA-based optical gates (from 60 mA to 150 mA) do not impact the received OSNR of decomposed hardware cubes too much. Meanwhile, when applying the NOS with larger port counts (>16) and a small SOA operation current (<100 mA), the received OSNR dramatically increases under a larger operation current.
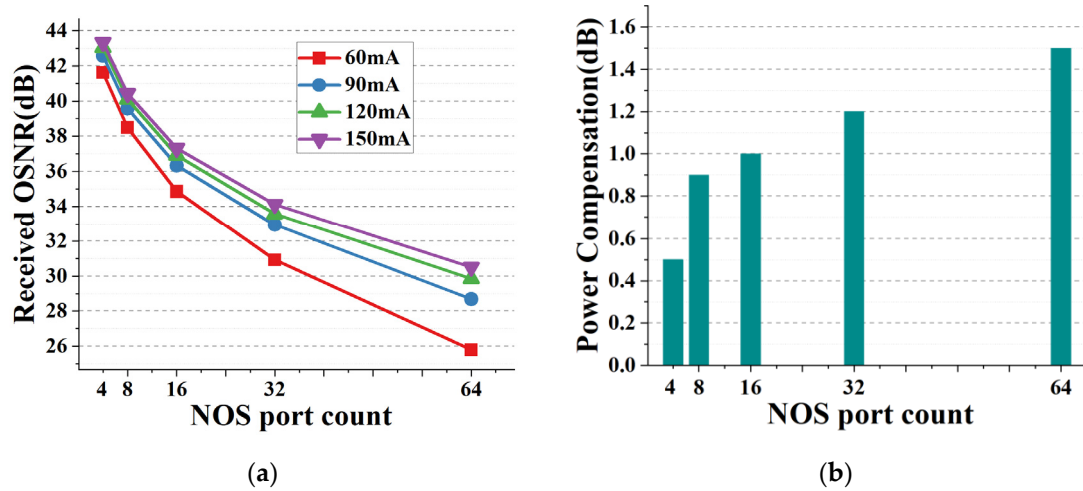

(a)

(b)

Figure 7. (**a**) Received ONSR of decomposed hardware under various operation currents; (**b**) Power compensation for none-error transmission ranging from four-port to 64-port NOS.

In order to guarantee the stable communication among decomposed hardware cubes, the required power compensation for none-error network transmission is investigated ranging from four-port NOS to 64-port NOS. In the experiment, the required power compensation is measured with a BER of 1×10^{-9}. As depicted in Figure 7b, the optical computing architecture with decomposed hardware requires a power compensation of 0.9 dB for the none-error transmission with an eight-port NOS. With a 64-port NOS, the required power compensation increases to 1.5 dB for none-error transmission. The reason is that, with the broadcast and select structure, there is much more power loss after optical signal splitters (18 dB extra loss under NOS port count of 64). When amplifying the optical signal at SOA-based optical gates, more noise is introduced under a lower input power. The experimental investigation shows that the proposed optical computing architecture with decomposed hardware is feasible and stable under a larger network scale.

5.2. Network Performance under Larger Network Scales

Due to the limitation of the hardware amount, it is not easy to experimentally evaluate the network performance of the implemented computing prototype with decomposed hardware cubes at larger scales. Thus, exploiting the experimental measurements (like the NOS switching time and processing time of processor/memory cubes), we numerically assess the network performance of optical computing architecture with decomposed hardware in this part. The discrete event simulator OMNeT++ is applied to model the NOS-based flat interconnect network and decomposed hardware cubes. Scaling the decomposed computing architecture, the network topology in [26] is applied in the assessment to interconnect hardware cubes. Four different network scales and NOS with a corresponding port count are considered in this evaluation: 64 hardware cubes (eight-port NOS), 256 hardware cubes (16-port NOS), 1024 hardware cubes (32-port NOS), and 4096 hardware cubes (64-port NOS). As analyzed in Section 3, the grouping number M is configured as 4, while each group includes two to 16 hardware cubes. Considering the network performance and cost, the TRX bandwidth in each hardware cube is configured as 40 Gb/s based on numerical investigation in [25]. Benefiting from the local memory in the processor cube and on-board resource in the memory cube, no packet loss can be measured in the experimental demonstration. Therefore, the end-to-end access latency is used as performance criterion in the assessment. To study the end-to-end access latency of the NOS-based computing architec-

ture with decomposed hardware, the Memory Cube Access Ratio (MCAR) is defined as the ratio of required memory size in memory cubes and the overall memory size required by processor cubes. The access instruction is sent from the processor cubes following existing computing structures [32]. Based on the spatial and temporal characteristics of data access, long-tailed Pareto distribution is applied to generate access instruction. Same as the experimental investigation, each network packet is configured as 512 bit-length. While a memory cube processes the data access instruction, the target cache line is first fetched to the processor cube. Then, the remaining data of the same data page (4096 bytes) are fetched to source the processor cube to avoid repeat remote memory cube access [33]. In the assessment, the network packets carrying target data are assigned with higher priority.

Figure 8a reports the end-to-end access latency of the target data line with different network scales and different locations of decomposed hardware cubes. There are two kinds of locations: the source processor cube and destination memory cube are located inside the same rack (intra-rack) and across different racks (inter-rack). It is illustrated that, with the MCAR of less than 0.9, the intra-rack and inter-rack locations perform similar end-to-end access latency with four different scales. With a network scale of 4096 decomposed hardware cubes, intra-rack access latency of 130.8 ns and inter-rack access latency of 240.9 ns are obtained. When the MCAR is larger than 0.9, the end-to-end access latency dramatically increases scaling the optical computing architecture with decomposed hardware cubes. However, while increasing decomposed hardware from 64 cubes to 4096 cubes, the intra-rack end-to-end access latency increases by 12.6% with the MCAR of 0.9 (148.5 ns for 4096 decomposed hardware cubes), compared with 15.8% more access latency for inter-rack interconnection (265.7 ns for 4096 decomposed hardware cubes).
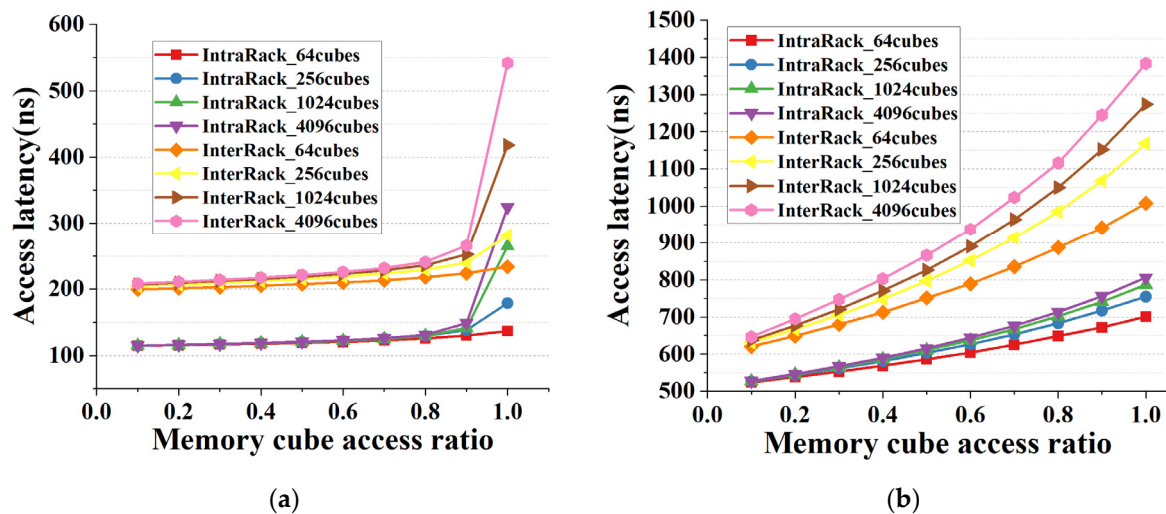

(a)   (b)

Figure 8. (**a**) End-to-end access latency of fetching (**a**) target data line and (**b**) data page with 40 Gb/s TRX bandwidth.

Besides the target data line access, the end-to-end access latency for data page access is also numerically investigated as shown in Figure 8b. Different with target data line access, the end-to-end latency for data page access increases with larger MCAR regardless of network scales. Scaling the amount of decomposed hardware from 64 cubes to 4096 cubes, the intra-rack end-to-end access latency increases by 11.1% under an MCAR of 0.9 (671.8 ns for 64-cube network and 755.8 ns for 4096-cube network). As for inter-rack interconnection, the end-to-end latency for data page access performs only a small increase of 7.4% under large network sizes (more than 256 cubes). It can be referred that the optical computing architecture keeps similar performance when increasing the port count of NOS and interconnecting more decomposed hardware cubes.

5.3. Discussion

Based on the above experimental and numerical assessments, it is shown that the NOS-based optical network can provide stable and fast switching interconnection among hardware cubes under various network scales. Compared with the microseconds access latency of current electrical network-based solutions in [16,17], the implemented decomposed optical computing architecture performs lower target data access latency of 148.5 ns under a 4096-cube network and an MCAR of 0.9. Meanwhile, the proposed NOS-based optical computing architecture can achieve a higher aggregated bandwidth (up to hundreds of gigabits) utilizing commercially available transceivers and Dense Wavelength Division Multiplexing (DWDM) technologies.

Compared with milliseconds switching time in an Optical Circuit Switch (OCS)-based network in [21] and microseconds switching time in a micro-ring resonator-based network in [23], the implemented NOS-based decomposed computing prototype can achieve nanoseconds switching time (43.4 ns), while keeping a stable interconnect link with none packet loss. These assessment results demonstrate the superiority of the proposed optical decomposed computing network based on NOS with respect to existing works.

6. Conclusions

We implement and experimentally investigate an optical computing prototype with decomposed hardware cubes for STN edge computing. In the proposed NOS-based flat terrestrial computing network, decomposed hardware cubes are interconnected, leveraging two parallel optical packet sending plane and control plane. Based on a four-cube computing prototype, the physical and network performance of the proposed computing architecture are experimentally assessed. Then, utilizing the experimentally measured parameters, the scalability of the decomposed computing network is numerically evaluated scaling up to 4096 hardware cubes. Finally, compared with the existing terrestrial computing networks, the superiority of the implemented optical computing network with decomposed hardware cubes is verified, and it is feasible to apply the proposed decomposed architecture as terrestrial infrastructures for STN edge computing.

It is shown in experimental and numerical assessments:

(1) In the physical assessment, the implemented computing prototype with decomposed hardware cubes achieves none-error packet transmission based on the power compensation of 0.5 dB. Minimal signal interference across the optical channel is ensured with larger than 60 dB ON/OFF ratio of SOA-based gates.

(2) For the network performance of the computing prototype with decomposed hardware cubes, an end-to-end access latency of 122.3 ns can be obtained in the experimental investigation, while there is zero packet loss after initial CDR procedure.

(3) When scaling the NOS port count to 64, the NOS-based interconnect network provides optical signals with 30.5 dB OSNR at the receiver part of decomposed hardware cubes, while requiring power compensation of 1.5 dB for none-error packet transmission. Under a network scale of 4096 decomposed hardware cubes, numerical studies report an end-to-end access latency of 148.5 ns inside the same rack with an MCAR of 0.9 and TRX bandwidth of 40 Gb/s.

Compared with current terrestrial computing networks, the proposed NOS-based flat decomposed computing network can provide high bandwidth optical interconnection among hardware cubes (even higher than hundreds of gigabits) and low access latency (tens nanoseconds). Benefiting from the parallel and flat structure, the performance is maintained when scaling the network (up to thousands of hardware cubes).

Due to the broadcast and select structure of NOS, the SOA-based gates receive much lower optical power input when further increasing the port count of NOS. This limitation may introduce more noise when amplifying the optical signals and degrade the performance of the NOS-based decomposed computing network. Meanwhile, if the target data are non-uniformly distributed among hardware cubes, there are more potential optical packet contentions, leading to higher access latency. Therefore, it is necessary to design

the corresponding data management algorithm and packet routing policy to minimize the packet contention and access latency.

Author Contributions: Conceptualization, X.G. and Y.J.; methodology, X.G.; software, X.G.; formal analysis, X.G., Y.Z., Y.J., S.W. and H.L.; resources, X.G. and Y.Z.; data curation, X.G.; writing—original draft preparation, X.G.; writing—review and editing, X.G.; visualization, X.G.; supervision, X.G.; project administration, X.G. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kim, T.; Kwak, J.; Choi, J.P. Satellite Edge Computing Architecture and Network Slice Scheduling for IoT Support. *IEEE Internet Things J.* **2021**, *8*, 1. [CrossRef]
2. Xie, R.; Tang, Q.; Wang, Q.; Liu, X.; Yu, F.R.; Huang, T. Satellite-Terrestrial Integrated Edge Computing Networks: Architecture, Challenges, and Open Issues. *IEEE Netw.* **2020**, *34*, 224–231. [CrossRef]
3. Wang, Y.; Yang, J.; Guo, X.; Qu, Z. Satellite edge computing for the internet of things in aerospace. *Sensors* **2019**, *19*, 4375. [CrossRef] [PubMed]
4. Li, C.; Zhang, Y.; Xie, R.; Hao, X.; Huang, T. Integrating edge computing into low earth orbit satellite networks: Architecture and prototype. *IEEE Access* **2021**, *9*, 39126–39138. [CrossRef]
5. Tang, Q.; Fei, Z.; Li, B.; Han, Z. Computation offloading in LEO satellite networks with hybrid cloud and edge computing. *IEEE Internet Things J.* **2021**, *8*, 9164–9177. [CrossRef]
6. Wang, Y.; Zhang, J.; Xing, Z.; Peng, W.; Liu, L. A Computation Offloading Strategy in Satellite Terrestrial Networks with Double Edge Computing. In Proceedings of the 2018 IEEE International Conference on Communication Systems (ICCS), Chengdu, China, 19–21 December 2018; pp. 450–455. [CrossRef]
7. Guo, J.; Chang, Z.; Wang, S.; Ding, H.; Feng, Y.; Mao, L.; Bao, Y. Who limits the resource efficiency of my datacenter: An analysis of alibaba datacenter traces. In Proceedings of the ACM 27th International Symposium on Quality of Service, Phoenix, AZ, USA, 24–25 June 2019; pp. 1–10.
8. Lin, R.; Cheng, Y.; Andrade, M.D.; Wosinska, L.; Chen, J. Disaggregated data centers: Challenges and trade-offs. *IEEE Commun. Mag.* **2020**, *58*, 20–26. [CrossRef]
9. Han, S.; Egi, N.; Panda, A.; Ratnasamy, S.; Shi, G.; Shenker, S. Network support for resource disaggregation in next generation datacenters. In Proceedings of the 12th ACM Workshop on Hot Topics in Netwworks (HotNets-XII), College Park, MD, USA, 21–22 November 2013.
10. Li, C.-S.; Franke, H.; Parris, C.; Abali, B.; Kesavan, M.; Chang, V. Composable architecture for rack scale big data computing. *Futur. Gener. Comp. Syst.* **2017**, *67*, 180–193. [CrossRef]
11. Bielski, M.; Syrigos, I.; Katrinis, K.; Syrivelis, D.; Reale, A.; Theodoropoulos, D.; Alachiotis, N.; Pnevmatikatos, D.; Pap, E.; Zervas, G.; et al. dReDBox: Materializing a full-stack rack-scale system prototype of a next-generation disaggregated datacenter. In Proceedings of the 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden, Germany, 19–23 March 2018; pp. 1093–1098. [CrossRef]
12. Alexoudi, T.; Terzenidis, N.; Pitris, S.; Moralis-Pegios, M.; Maniotis, P.; Vagionas, C.; Mitsolidou, C.; Mourgias-Alexandris, G.; Kanellos, G.T.; Miliou, A.; et al. Optics in Computing: From Photonic Network-on-Chip to Chip-to-Chip Interconnects and Disintegrated Architectures. *J. Light. Technol.* **2019**, *37*, 363–379. [CrossRef]
13. Pfandzelter, T.; Bermbach, D. QoS-Aware Resource Placement for LEO Satellite Edge Computing. In Proceedings of the 6th International Conference on Fog and Edge Computing (ICFEC), Messina, Italy, 16–19 May 2022; pp. 66–72.
14. Tong, M.; Wang, X.; Li, S.; Peng, L. Joint Offloading Decision and Resource Allocation in Mobile Edge Computing-Enabled Satellite-Terrestrial Network. *Symmetry* **2022**, *14*, 564. [CrossRef]
15. Gao, P.X. Network requirements for resource disaggregation. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI), Savannah, GA, USA, 2–4 November 2016; pp. 249–264.
16. Lim, K.; Turner, Y.; Santos, J.R.; AuYoung, A.; Chang, J.; Ranganathan, P.; Wenisch, T.F. System-level implications of disaggregated memory. In Proceedings of the IEEE International Symposium on High-Performance Comp Architecture, New Orleans, LA, USA, 25–29 February 2012; pp. 1–12.

17.	Shan, Y.; Huang, Y.; Chen, Y.; Zhang, Y. LegoOS: A Disseminated, Distributed OS for Hardware Resource Disaggregation. In Proceedings of the 13th USENIX Symposium on Operating Systems Design and Implementation (OSDI '18), Carlsbad, CA, USA, 8–10 October 2018; pp. 69–87.

18.	Abrahamse, R.; Hadnagy, A.; Al-Ars, Z. Memory-Disaggregated In-Memory Object Store Framework for Big Data Applications. *arXiv* **2022**, arXiv:2204.12889.

19.	Guo, X.; Xue, X.; Yan, F.; Pan, B.; Exarchakos, G.; Calabretta, N. DACON: A reconfigurable application-centric optical network for disaggregated data center infrastructures [Invited]. *J. Opt. Commun. Netw.* **2022**, *14*, A69–A80. [CrossRef]

20.	Ali, H.M.M.; El-Gorashi, T.E.H.; Lawey, A.Q.; Elmirghani, J.M.H. Future energy efficient data centers with disaggregated servers. *J. Light. Technol.* **2017**, *35*, 5361–5380.

21.	Saljoghei, A.; Yuan, H.; Mishra, V.; Enrico, M.; Parsons, N.; Kochis, C.; De Dobbelaere, P.; Theodoropoulos, D.; Pnevmatikatos, D.; Syrivelis, D.; et al. MCF-SMF Hybrid Low-Latency Circuit-Switched Optical Network for Disaggregated Data Centers. *J. Light. Technol* **2019**, *37*, 4017–4029. [CrossRef]

22.	Zhu, Z.; Di Guglielmo, G.; Cheng, Q.; Glick, M.; Kwon, J.; Guan, H.; Carloni, L.P.; Bergman, K. Photonic Switched Optically Connected Memory: An Approach to Address Memory Challenges in Deep Learning. *J. Light. Technol* **2020**, *38*, 2815–2825. [CrossRef]

23.	Gonzalez, J.; Palma, M.G.; Hattink, M.; Rubio-Noriega, R.; Orosa, L.; Mutlu, O.; Bergman, K.; Azevedo, R. Optically connected memory for disaggregated data centers. *J. Parallel Distrib. Comput.* **2022**, *163*, 300–312. [CrossRef]

24.	Ballani, H.; Costa, P.; Behrendt, R.; Cletheroe, D.; Haller, I.; Jozwik, K.; Karinou, F.; Lange, S.; Shi, K.; Thomsen, B.; et al. Sirius: A flat datacenter network with nanosecond optical switching. In Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM'20), Association for Computing Machinery, New York, NY, USA, 10–14 August 2020; pp. 782–797.

25.	Guo, X.; Yan, F.; Wang, J.; Exarchakos, G.; Peng, Y.; Xue, X.; Pan, B.; Calabretta, N. RDON: A rack-scale disaggregated data center network based on a distributed fast optical switch. *J. Opt. Commun. Netw.* **2020**, *12*, 251–263. [CrossRef]

26.	Yan, F.; Miao, W.; Raz, O.; Calabretta, N. OPSquare: A flat DCN architecture based on flow-controlled optical packet switches. *J. Opt. Commun. Netw.* **2017**, *9*, 291–303. [CrossRef]

27.	Miao, W.; Luo, J.; Lucente, S.D.; Dorren, H.; Calabretta, N. Novel flat datacenter network architecture based on scalable and flow controlled optical switch system. *Opt. Express* **2014**, *22*, 2465–2472. [CrossRef] [PubMed]

28.	Drepper, U. What Every Programmer Should Know about Memory. Available online: http://people.redhat.com/drepper/cpumemory.pdf (accessed on 9 February 2015).

29.	Xilinx. UltraScale Architecture and Product Data Sheet: Overview. Available online: https://www.xilinx.com/products/silicon-devices/fpga/virtex-ultrascale.html (accessed on 20 May 2022).

30.	Xilinx. VC709 Evaluation Board for the Virtex-7 FPGA. Available online: https://www.xilinx.com/products/boards-and-kits/dk-v7-vc709-g.html (accessed on 11 March 2019).

31.	Xilinx. IBERT for UltraScale/UltraScale+ GTH Transceivers. Available online: https://www.xilinx.com/products/intellectual-property/ibert_ultrascale_gth.html (accessed on 4 February 2021).

32.	Hennessy, J.L.; Patterson, D.A. *Computer Architecture: A Quantitative Approach*, 6th ed.; Elsevier: Amsterdam, The Netherlands, 2018.

33.	Matani, D.; Shah, K.; Mitra, A. An O(1) Algorithm for Implementing the lfu CACHE eviction Scheme. *arXiv* **2021**, arXiv:2110.11602.