JOURNAL OF
Optical Communications
and Networking

Flexible low-latency metro-access converged network architecture based on optical time slice switching

Jialong Li, Nan Hua, Zhizhen Zhong, Yufang Yu, Xiaoping Zheng,* and Bingkun Zhou

Beijing National Research Center for Information Science and Technology (BNRist), Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
**Corresponding author: xpzheng@mail.tsinghua.edu.cn*

Received 16 July 2019; revised 23 October 2019; accepted 25 October 2019; published 21 November 2019 (Doc. ID 372926)

The "pay-as-you-grow" cloud computing model has become popular for today's enterprises. Cloud computing not only frees end users from complex operations, but also allows higher resource utilization, lower investment, and increased energy efficiency. However, with some emerging technologies, cloud computing is unable to meet the required latency level, especially for delay-sensitive services such as 5G communications, live streaming, and online gaming. In this context, edge computing is regarded as a promising technology that can provide low-latency connections in the near future. Unlike cloud computing, which provides service in a centralized mode, edge computing deploys micro data centers (MDCs) at the edge of the network to provide rapid-response service. However, due to the limited computing and storage capacity of a single MDC, a user may not be able to access the resources from the closest MDC during peak traffic periods. Under such circumstances, the user is served through another MDC or a remote cloud data center. The data are processed in an optical line terminal and then transmitted via the metro network, which significantly increases the latency. In this study, we introduce a flexible low-latency metro-access converged network architecture based on optical time slice switching (OTSS) to address the latency problem. By leveraging the transparent connections of the OTSS in this novel architecture, data can be transmitted through the MDCs without requiring extra processing time. The simulation results demonstrate that our proposed architecture can provide lower-latency connections under a range of conditions, with a negligible decrease in network throughput, compared with an existing representative architecture. Additionally, we conducted experiments to validate the feasibility of our approach.

https://doi.org/10.1364/JOCN.11.000624

1. INTRODUCTION

The growing popularity of emerging technologies such as 5G communications, live streaming, and health-related services has led to an increased demand from subscribers for low latency in computing services, due to the significant impact of latency on the quality of services (QoS) and quality of experiences (QoE). For example, in 5G communications, the fronthaul and backhaul networks require very low latency to realize higher data transmission rates [1]. Ultra-reliable and low-latency communication (uRLLC) is one of the three application scenarios proposed by the 3rd Generation Partnership Project (3GPP) [2], which aims to provide less than 1 ms end-to-end latency to subscribers [3]. However, achieving this latency may be difficult, considering the long distances that may be between the users and the cloud data center. Moreover, the processing time required in the metro network may be non-negligible [4,5]. Another service with extremely strict latency requirements is virtual reality (VR). Because of the huge amount of processing data and requirements such as a light-weight VR machine on the user side, VR service providers tend to deploy data processing units at remote sites such as cloud data centers, which significantly increases the latency. However, dissonance occurs if the motion-to-photon latency is longer than 15–20 ms while experiencing VR scenarios [6].

Furthermore, users may experience high latency with the cloud computing model with services such as infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) [7], where data storage and computing resources are typically supported by cloud data centers consisting of thousands of co-located servers. Therefore, edge computing, which is regarded as a key technology in the Internet of Things (IoT) [8], has been introduced to satisfy the required levels of latency. The main idea of edge computing is deploying storage and computing resources at the edge of the network, enabling

rapid processing of user requests. The micro data centers (MDCs) deployed at the edge of the network are responsible for providing service, as compared to cloud data centers in the cloud computing model [9,10]. However, there are still some challenges with the practical implementation of edge computing [11,12]. First, due to the limited storage capacity of a single MDC, users may require services through distant MDCs during traffic peaks; we will refer to these as remote MDCs in this study. The situation is exacerbated by the tidal effect, which is caused by users moving between business and residential areas [13]. The propagation delay, which is a function of the distance between the users and the data center, and the processing time in the metro network may introduce excessive latency. Second, the required resources are not always stored in the closest MDC for some subscribers. In such cases, the user requires access to the resources stored in remote MDCs. Intelligent resource provisioning may alleviate the latency to some extent; however, it cannot entirely solve the problem because an accurate prediction of resource consumption is difficult [14]. Hence, these two challenges should be considered carefully to ensure the practical operation of edge computing.

Recently, a novel all-optical switching technology called optical time slice switching (OTSS) has been proposed to achieve transparent (i.e., all-optical) connections [15]. In this study, we have proposed a new architecture where data can be transmitted in a metro area without requiring extra processing time by leveraging the transparent connections of OTSS. The simulation results show that this novel architecture can achieve lower-latency connections compared with optical access networks based on a passive optical network (PON). In addition, we conducted experiments to further validate the feasibility of the proposed architecture.

The rest of the paper is organized as follows. In Section 2, we review previous work on edge computing and outline the challenges involved in its actual deployment. In Section 3, we present the flexible low-latency metro-access converged network approach based on OTSS. In Section 4, we discuss the compatibility of our proposed approach with existing network architectures. In Sections 5 and 6, we present the details of the simulation evaluation and experimental validation, respectively. Finally, in Section 7, we present the conclusions drawn from this study.

2. EDGE COMPUTING SCENARIOS IN METRO-ACCESS NETWORKS

Existing access networks widely use PON technology, due to its passive features and cost efficiency. Specifically, Gigabit PON (GPON) and Ethernet PON (EPON), which have been developed by ITU-T and IEEE, respectively, are the two most widely used PON technologies. In early 2010, the Full Service Access Network (FSAN) group initiated the study of next-generation optical access technologies, as part of the NG-PON2 project. Among the diverse candidate access technologies considered, FSAN selected the time- and wavelength-division multiplexing PON (TWDM-PON) technology in 2012 [16]. TWDM-PON attracted the interest of researchers after its specifications were released [17,18]. For example, a TWDM-PON prototype system was developed

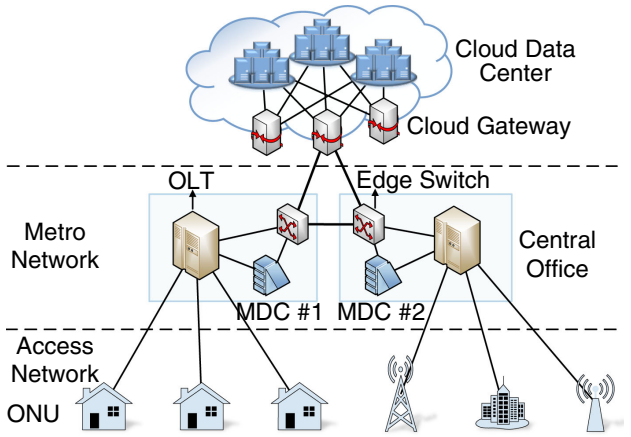

Fig. 1. Edge computing scenarios in metro-access networks.

and its wavelength tuning cost and load balancing capacity were tested in [19]. Regarding commercial viability, a 10 Gb/s tunable receiver was developed [20], which has been successfully verified by several companies and recently deployed by several service providers. Additionally, transceivers based on the specifications of NG-PON2 have been commercially produced [21]. Given the promise of TWDM-PON as a future technology, we have chosen it as the baseline for our new approach. We later compare the performance of our new approach with that of conventional TWDM-PON.

Figure 1 illustrates the edge computing scenario in metro-access networks. The optical network units (ONUs) and optical line terminals (OLTs) are the two main components of the TWDM-PON architecture in access networks. The ONUs provide network service to subscribers through the user interface, while the OLTs are responsible for data aggregation and transmission. Colorless ONUs can use any wavelength because tunable transceivers are implemented. Passive power splitters lie between the ONUs and the OLT. The OLTs and MDCs located in the central office (CO) and the metro network are responsible for the communications between the COs and the cloud data center. The MDCs contain storage and computing resources that can simultaneously be shared by multiple users. They are similar to cloud data centers, except the resources are significantly closer to the users than in cloud computing, thus enabling the users to experience low-latency services, which is very important for some delay-sensitive applications.

Many researchers have realized the benefits of implementing edge computing facilities by leveraging the PON architecture. For instance, a PON-enabled programmable access and edge cloud architecture was proposed in [22], where the CO was re-architected as an MDC. In this architecture, users can experience low-latency connections when the service is provided by the closest MDC. However, a processing time may be introduced by the metro network when a user needs to connect to a remote MDC or the cloud data centers. In [23], the authors presented a localization scheme to integrate the MDCs and OLTs. Many localization services can be processed by an MDC located in an OLT. The results demonstrated the advantages

of deploying MDCs at the edge of the network, and the proposed scheme could realize a higher data transmission rate and reduced packet loss rate.

3. OTSS-ASSISTED FLEXIBLE LOW-LATENCY METRO-ACCESS CONVERGED NETWORK ARCHITECTURE

A. Introduction of OTSS

Figure 2 depicts the working mechanism of OTSS in optical access networks. A time-domain partition of a wavelength provides an optical transmission channel. The channel is divided into repetitive OTSS frames in the time domain with a period of T_F. Each OTSS frame can accommodate one or more optical time slices of variable lengths. As shown in the figure, the data from ONU #1 are transmitted through the repetitive optical time slice #A. When the optical time slice starts, the timing is called the switching point, and the high-speed optical switch operates according to the control signal from the switch controller. In our case, when the optical time slice #A starts, the high-speed optical switch directs it to the desired output port, OLT #2 in the figure. High-precision synchronized time technology guarantees the correct operation of the high-speed optical switch.

Generally, optical circuit switching (OCS), optical packet switching (OPS), and optical burst switching (OBS) are the three main optical switching technologies [24]. Among these, OCS enables optical bypass and provides connection at the wavelength level, leading to inefficient network utilization. OPS achieves finer granularity (packet level) compared with OCS. However, the immaturity of optical buffering is a major shortcoming of OPS. Meanwhile, OBS is an intermediary approach between OCS and OPS, which can potentially realize sub-wavelength granularity without requiring optical buffering [25]. However, OBS has not been shown to achieve reliable data transfer, due to the high potential for lost bursts [26]. As opposed to these strategies, OTSS uses time synchronization technology to precisely operate the high-speed optical switch periodically to realize sub-wavelength granularity. Therefore, an OTSS network does not require optical buffers and achieves reliable transmission.

B. Metro-Access Network Architecture Based on OTSS

Figure 3 illustrates a metro-access network based on OTSS. Two COs, namely, CO #1 and CO #2, are depicted in Fig. 3. For simplicity, we refer to CO #1 and CO #2 by their corresponding access networks PON #1 and PON #2, respectively. Three upstream data flows, labeled as A and B in PON #1, and C in PON #2, are generated by their corresponding ONUs. Data flow A is transmitted in the traditional manner, while the transmission of B and C is based on OTSS. Similarly, data flow A is encapsulated in the PON frame, while B and C are encapsulated in the OTSS frame.

In the edge computing scenario, the latency would be low if the ONU were served by the closest MDC. For example, if an ONU from PON #1 requested a service, then MDC #1, which is located in CO #1, would be the best choice. We use data flow A in PON #1 to represent this scenario. However, due to the limited storage and computing capacity of MDC #1, the ONU may need to connect to a remote MDC, such as MDC #2 or a remote cloud data center (the cloud data center is not shown in the figure). First, we consider how this scenario would be implemented in a more conventional PON-based metro architecture. The data must first traverse to CO #1 and then be transmitted to MDC #2 through the edge switches implemented in the CO, potentially resulting in an excessive processing delay. We use the combination of data flows A and D to represent this case. Data flow D is the transmission between COs #1 and #2 after data flow A arrives at CO #1, and is encapsulated in the optical transport network (OTN) or packet transport network (PTN) frame, which is transmitted through the metro network.

We next consider the situation when the metro-access converged architecture is based on OTSS; i.e., we refer to this proposed architecture as TS-TWDM-PON. Unlike the network architecture described above, cross-PON switches (CPSs) are deployed in TS-TWDM-PON. We adopt a centralized control mode, with the central controller (implemented in one of the COs) responsible for the correct operation of the CPSs. We have investigated the controller problem in our previous work [27]. When using the OTSS technology, data can directly

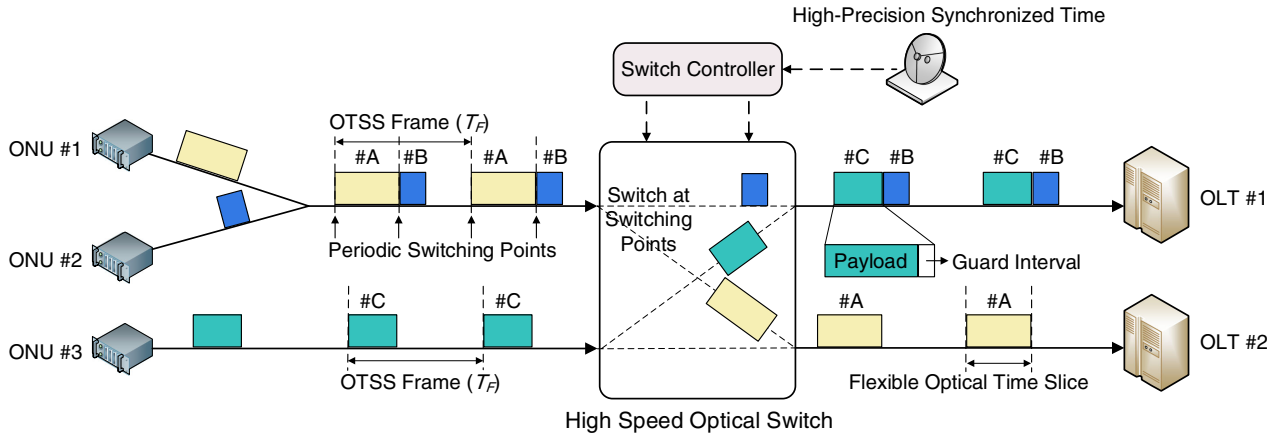

Fig. 2. Working mechanism of OTSS in optical access networks.

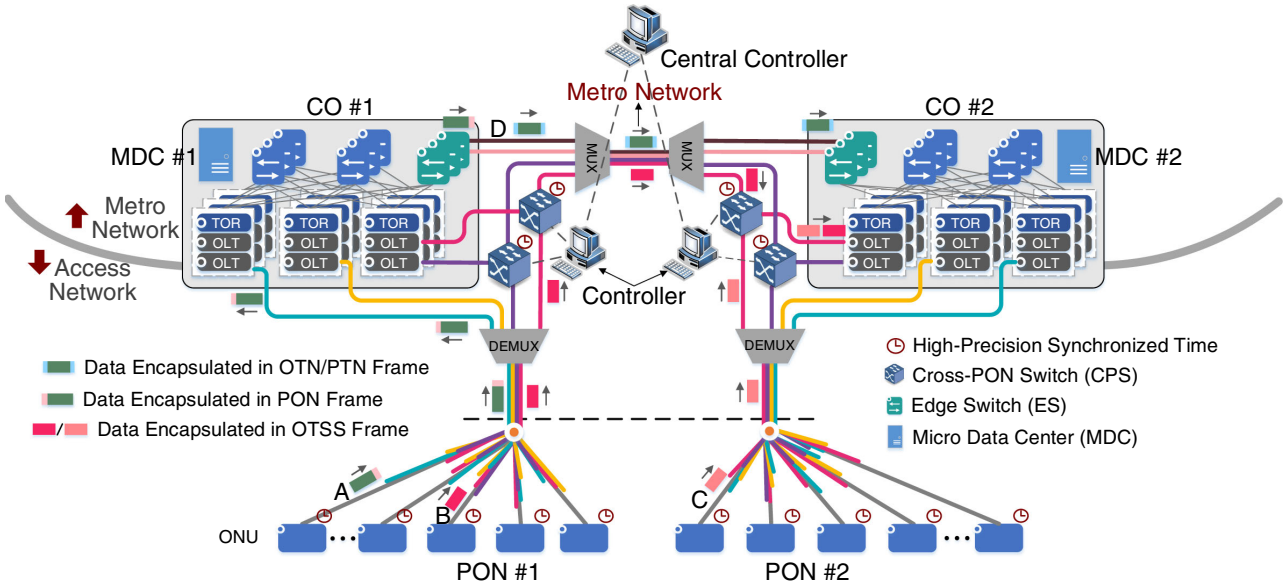

Fig. 3. Metro-access converged network based on OTSS.

be sliced and all-optically switched by the CPSs without passing through the edge switches. Thus, processing latency at the edge switches is eliminated. For example, the destination of data flow B is MDC #2. This data flow is encapsulated in the OTSS frame and passes through the metro-access converged network all-optically, without requiring processing time in a switch. Similarly, data flow C is also encapsulated in the OTSS frame, and its destination is the closest edge computing node, i.e., MDC #2. Data flows B and C can be transmitted to MDC #2 without collision, by utilizing time synchronization technology.

C. Bandwidth Assignment Algorithm

In a conventional PON architecture, the OLT grants access to each ONU in the upstream channel. An ONU reports the number of required time slots, and the OLT grants the time slots based on this information. However, in the TS-TWDM-PON approach, the bandwidth assignment is more complicated. Because of the implementation of the OTSS, the channel is divided into repetitive frames. Each OTSS frame is further divided into many small slices, as shown in Fig. 4. A slice represents the finest partitioning in the time domain; we refer to it as the minimum optical time slice. The bandwidth of the minimal optical time slice is thus the bandwidth granularity of the OTSS. A request is accommodated by an optical time slice, which consists of several contiguous minimum optical time slices. For example, assume that the lengths of the OTSS frame and minimum optical time slice are set to 100 μs and 1 μs, respectively, the wavelength capacity is set to 10 Gb/s, and the required bandwidth of a request is 1 Gb/s. Then, the corresponding bandwidth of a minimum optical time slice, i.e., the bandwidth granularity of the OTSS, is 100 Mb/s, and the 1 Gb/s request requires 10 contiguous minimum optical time slices. Note that a finer bandwidth granularity requires stricter time synchronization precision in the network.

Furthermore, note that the start and end positions of the contiguous optical time slices at different links are different because of the propagation time [28]. For example, assume that a request occupies the contiguous optical time slices from positions 1 to 10 in the OTSS frames at Link A. Assume that this request is transmitted from Link A to Link B and the propagation delay between the two links is 30 μs. In this case, the start and end positions of the contiguous optical time slices are 31 and 40 at Link B, respectively. Therefore, we must ensure that the contiguous optical time slices from position 31 to 40 are available for this request at Link B.

We have developed a wavelength and time slice allocation (WTSA) algorithm for the TS-TWDM-PON. Algorithm 1 describes the WTSA in detail. Lines 2 and 3 calculate the number of contiguous minimum optical time slices for a request, which is represented by N in the algorithm. In Lines 4 to 10, we find a wavelength with N contiguous minimum optical time slices at all links, while considering the propagation delay. If the wavelength is already heavily used and fails to support the request, the ONU can tune to another wavelength to try to find available optical slices because the ONU is equipped with tunable transceivers.

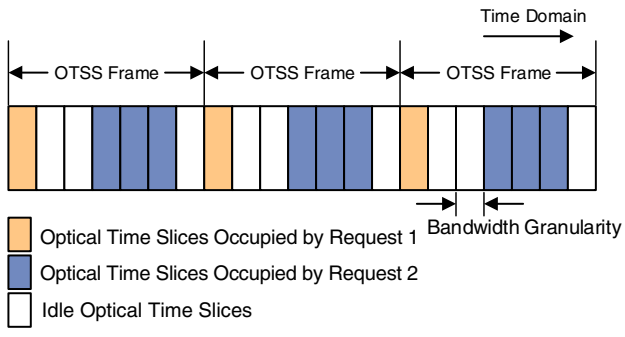

Fig. 4. Optical time slice assignment for different requests.

Algorithm 1. Wavelength and Time Slice Allocation Algorithm

Require:
$\mathcal{R} \leftarrow$ the set consisting of requests;
$\mathcal{W} \leftarrow$ the wavelength set of each PON system;
$B_i \leftarrow$ required bandwidth of request i (unit: Mb/s);
$T_F \leftarrow$ the time duration of an OTSS frame (unit: μs);
$T_m \leftarrow$ the time duration of a minimum optical time slice (unit: μs);
$C \leftarrow$ capacity of a single wavelength (unit: Mb/s);
1: select a request i from $\mathcal{R}$;
2: calculate the required number of continuous minimum optical time slices N_i;
3: $N_i \leftarrow \lceil B_i T_F / C T_{mo} \rceil$ for request i;
4: **if** the state of request i is arriving **then**
5: **for** wavelength j in $\mathcal{W}$ **do**
6: **if** N_i continuous minimum optical time slices are available in wavelength j at all links **then**
7: use this wavelength to accommodate the request i;
8: mark the N_i optical time slices occupied at all links;
9: **break**;
10: **end if**
11: **end for**
12: **else**
13: the state of request i is leaving;
14: release the N_i optical time slices occupied at all links;
15: **end if**
16: **if** all wavelengths fail to accommodate request i **then**
17: block this request;
18: **end if**
19: calculate the total latency of request i;
20: return to step 1 until $\mathcal{R}$ is empty;
21: calculate the average latency of all requests.

Herein, we analyze the time complexity of the WTSA algorithm. We adopt the First Fit algorithm to search for the available optical time slices for a request. The computational complexity of WTSA is $O(WN^2)$, where W and N represent the number of wavelengths and the number of minimum optical time slices in an OTSS frame (the length of an OTSS frame divided by the length of a minimum optical time slice), respectively.

4. COMPATIBILITY OF TS-TWDM-PON

To evaluate the compatibility of TS-TWDM-PON, we need to consider the functioning of existing PON architectures and the feasibility of TS-TWDM-PON. As mentioned earlier, GPON and EPON are the two most widely used technologies. Hence, in this section, we contrast our proposed architecture with that of GPON and EPON.

A. Time Synchronization Technology

In the upstream, the ONU transmits data only during the time slots assigned by the OLT, which is called time-division multiple access (TDMA). To achieve this, the OLT knows the precise arrival time of the data from each ONU and globally manages the time slot assignment to avoid collision. The

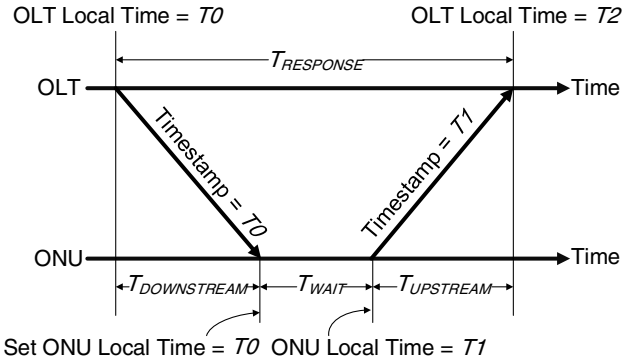

Fig. 5. Process of distance measurement [29].

PON system uses the round-trip time (RTT) measurement to achieve time synchronization between the OLT and ONUs. Figure 5 depicts the distance measurement process [29]. Note that the $T_{\text{DOWNSTREAM}}$ and T_{UPSTREAM} times represent the downstream and upstream propagation delay, respectively; and the RTT between the OLT and the ONUs is calculated as follows:

$$\begin{aligned} \text{RTT} &= T_{\text{DOWNSTREAM}} + T_{\text{UPSTREAM}} \\ &= T2 - T1 \end{aligned} \tag{1}$$

However, the distance measurement is applicable only in a star topology, i.e., a point-to-multipoint architecture, while the TS-TWDM-PON is a multipoint-to-multipoint topology because there are multiple COs and ONUs. In our architecture, the time of the CPSs must be synchronized with the ONUs belonging to different PONs. Therefore, the distance measurement cannot be applied in TS-TWDM-PON. Furthermore, this can achieve only microsecond-level time synchronization, and hence, it is difficult to meet the requirements of the fast-growing wireless technologies [30]. Fortunately, we can use a mature sub-microsecond network time synchronization technology to realize cross-PON synchronization [31], such as IEEE 1588v2 [32], which has achieved sub-microsecond-level accuracy in deployed networks and an accuracy of approximately 50 ns in testbeds [33]. Furthermore, our lab has proposed a high-reliability sub-nanosecond network time synchronization method, enabling a 50-ns-level accuracy that is applicable in metro and backbone networks [34].

Asymmetry is certainly a problem that needs to be considered in TS-TWDM-PON. Actually, 1588v2 does not include specifications on asymmetry, although according to the recommendation, the hardware should support asymmetry-compensation solutions. Asymmetry is introduced mainly by two factors: propagation delay and transmission-delay variation. Propagation-delay asymmetry is caused by different fiber lengths or wavelengths used in two transmission directions. This kind of asymmetry is relatively static and relatively easy to determine [35]. On the other hand, transmission-delay variation is induced by fluctuations in environmental temperature. This type of delay introduces time-varying synchronization errors and can be compensated for in real time, resulting in

only a slight residual impact on sub-nanosecond-level time synchronization systems [36,37].

As a larger number of COs and longer distances deteriorate the synchronization accuracy, wider guard intervals are required between different optical time slices, which reduces the effective bandwidth. Under the sub-microsecond-level accuracy provided by 1588v2, the sub-microsecond-level guard intervals of the OTSS optical slices from different ports can be considered as a feasible value that ensures the correct switching operation of the switches. In our experiment, the guard interval was set to 100 ns (this will be discussed further in Section 6).

Time synchronization technology enables collaborative time slot assignment within different PONs. In addition, 10-ns-level optical switches [38] can be implemented in the CPSs to achieve fine-grained optical time slice switching. All CPSs can be implemented inside the COs to obtain power supply, and an important characteristic of our new approach is that it does not change the passive features of the PON architecture.

B. Frame Structure

The data from different ONUs are encapsulated in the upstream frame and transmitted to the OLT. In the existing commercial EPON and GPON, the length of the upstream/downstream frame is fixed. Figures 6 and 7 illustrate the upstream frame structure in EPON and GPON, respectively. The lengths of the upstream/downstream frames in EPON and GPON are 2 ms and 125 μs, respectively. In EPON, the Ethernet frames are filled in the upstream frame payload. An Ethernet frame comprises a preamble, variable payload, and frame check sequence (FCS). The preamble has been modified to include a logical link identifier (LLID), which is used by the OLT to distinguish different ONUs. In general, guard intervals are required between different time slots to avoid data collision. GPON leverages GPON encapsulation mode (GEM), wherein the Ethernet packets are first mapped into GEM frames, which are filled in the GPON upstream payload. As for the downstream frames of EPON and GPON, the OLT broadcasts all the packets from the metro network, while each ONU extracts only its corresponding data and ignores the other packets.

The OTSS frame design is shown in Fig. 8. The length of the OTSS frame is variable and typically ranges from several hundred μs to several hundred ms. A request occupies one or more contiguous minimum optical time slices, typically ranging from one to several μs. A burst overhead is required for

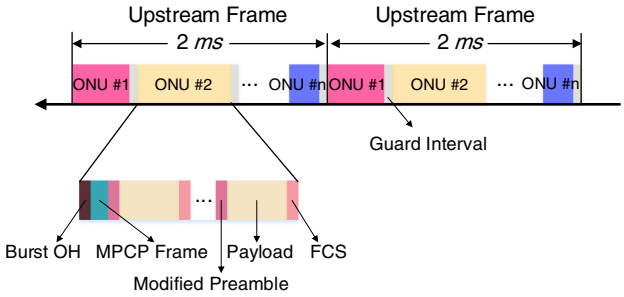

Fig. 6. EPON upstream frame. OH, overhead; FCS, frame check sequence; MPCP, multi-point control protocol.

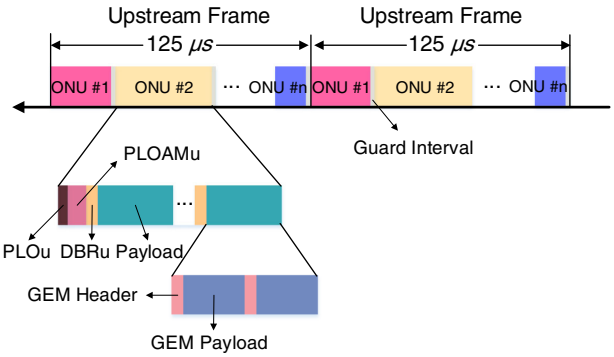

Fig. 7. GPON upstream frame. PLOu, physical layer overhead upstream; OAM, operation, administration, and maintenance; PLOAMu, physical layer OAM upstream; DBRu, dynamic bandwidth report upstream; GEM, GPON encapsulation mode.

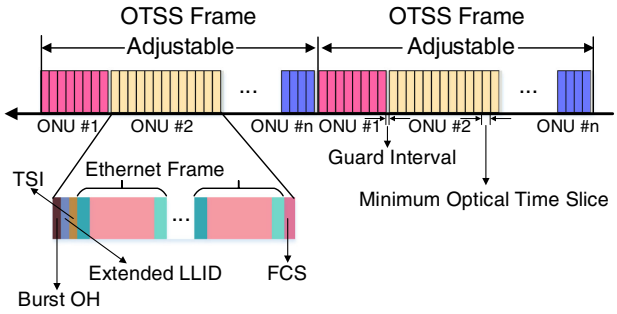

Fig. 8. OTSS frame. OH, overhead; LLID, logical link identifier; TSI, time slice information.

burst-mode receiving, and each ONU owns a unique extended LLID (ELLID) in the network. Compared with the LLID applied in EPON, the ELLID consists of both a PON ID and an ONU ID. The optical time slice information (TSI) is designed for message interaction between the OLT and ONUs.

C. Upstream Signaling Design

As mentioned earlier, in the conventional PON architecture, the OLT grants access to each ONU in the upstream channel: before an ONU transmits the upstream data, the OLT sends a control message to this ONU, which includes the time slot information, and the ONU transmits its data according to the control message. In TS-TWDM-PON, the ONU transmits data not only to its associated OLT, but possibly also directly to a remote OLT.

Herein, we propose an upstream signaling design, as shown in Fig. 9. We first consider a case wherein data are transmitted within the same PON. In this case, the data need not traverse the metro network. The following procedure is involved:

1. The ONU sends a services request (REQU); the OLT determines that the request can be served by the local MDC.
2. The OLT sends TSI_REPORT, reporting the optical time slice usage information to the controller.

3. The OLT forwards GRANT_INNER containing the time slot information, instructing the ONU to transmit data.
4. Transmission begins.

The second case involves transmitting data between different PONs. In this case, the data must traverse the metro network, and hence, a central controller is required to coordinate the optical time slice assignment. The following procedure is involved:

a. The ONU sends a services request (REQU); the OLT determines that the request must be served by a remote MDC.
b. The OLT sends REQU_REPORT to the controller.
c. The controller sends SWITCHING_SIGNAL to the remote OLT.
d. The controller sends SWITCHING_SIGNAL to the OLT.
e. The OLT forwards GRANT_INTER containing the time slot information, instructing the ONU to transmit data.
f. Transmission begins.

5. SIMULATION EVALUATION

A. Simulation Setup

We developed a simulation platform based on Python to evaluate the performances of our proposed architecture. Figures 10(a) and 10(b) demonstrate the simulation topologies based on TWDM-PON and TS-TWDM-PON, respectively. The metro network is a ring topology with bidirectional transmission. The number of COs is four, and four OLTs are implemented in each CO. The number of ONUs is 32 in each PON system. Note that the MDCs deployed in the COs have not been depicted in the figure. Each CO is equipped with four wavelengths, and the capacity of a single wavelength is set to 10 Gb/s. The distance between each adjacent CO and the distance between the ONU and CO is set to 10 km.

In our simulation, we consider processing delay, propagation delay, and buffer delay. The total delay is the sum of the three types of delay. Processing delay is caused by electronic processing in the edge switches, including queuing in the routers and multiplex/demultiplex. The one-hop processing delay is related to the data transmission rate according to [5]. We use a linear model for simplicity. Based on the traffic model we use, 5 Erlang and 100 Erlang corresponding to around 475 Mb/s and 9.5 Gb/s per wavelength, respectively. The one-hop processing times of 5 Erlang and 100 Erlang are 300 μs and 6 ms, respectively. Propagation delay is caused by signal propagation. Generally, an optical signal propagates at a speed of 200 km/ms in fiber. Buffer delay refers to the waiting time for the packets in the ONU's buffer before transmission begins; i.e., when a packet arrives at the ONU, it cannot be transmitted immediately; it is stored in the buffer until the start of its assigned time slots.

In the topology based on TWDM-PON depicted in Fig. 10(a), the inter-CO data flows are forwarded through the electrical edge switches implemented at the COs, whereas, for the proposed TS-TWDM-PON approach shown in Fig. 10(b), they are all-optically switched at the CPSs deployed at the COs.

B. Network Delay Performances Under Uniform Traffic Load

To realize a uniform traffic load, we randomly selected one ONU and one CO as the source and destination of a request, respectively. The load on a CO is represented by the total bandwidth of requests whose destination is the CO. The load of the four COs is approximately the same. The traffic requests arrive according to a Poisson process. The bandwidth of a request is uniformly distributed from 50 Mb/s to 3000 Mb/s with a granularity of 50 Mb/s. The length of the OTSS frame is set to 1 ms. The difference between TS-TWDM-PON-I and TS-TWDM-PON-II is the bandwidth granularity, which is 5 μs and 10 μs, respectively. Considering that the bandwidth of a single wavelength is 10 Gb/s, the bandwidth granularities for TS-TWDM-PON-I and TS-TWDM-PON-II are

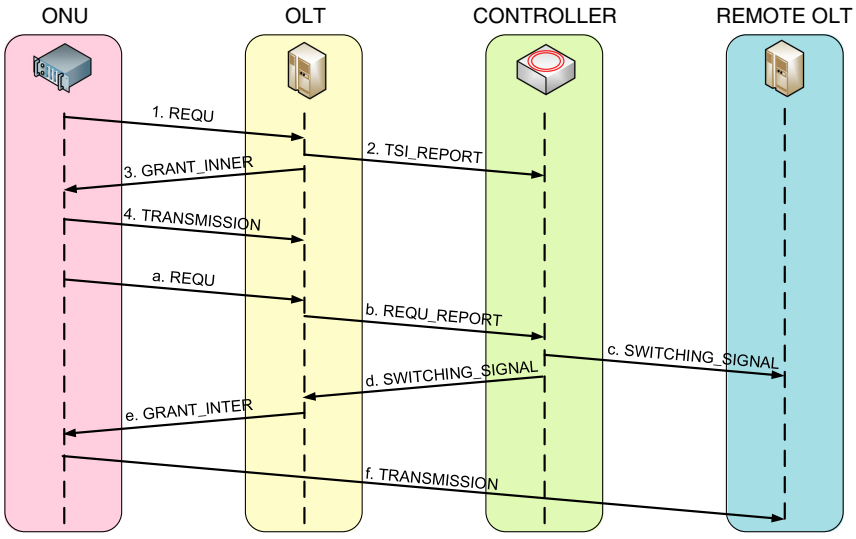

Fig. 9. Upstream signaling design for TS-TWDM-PON.

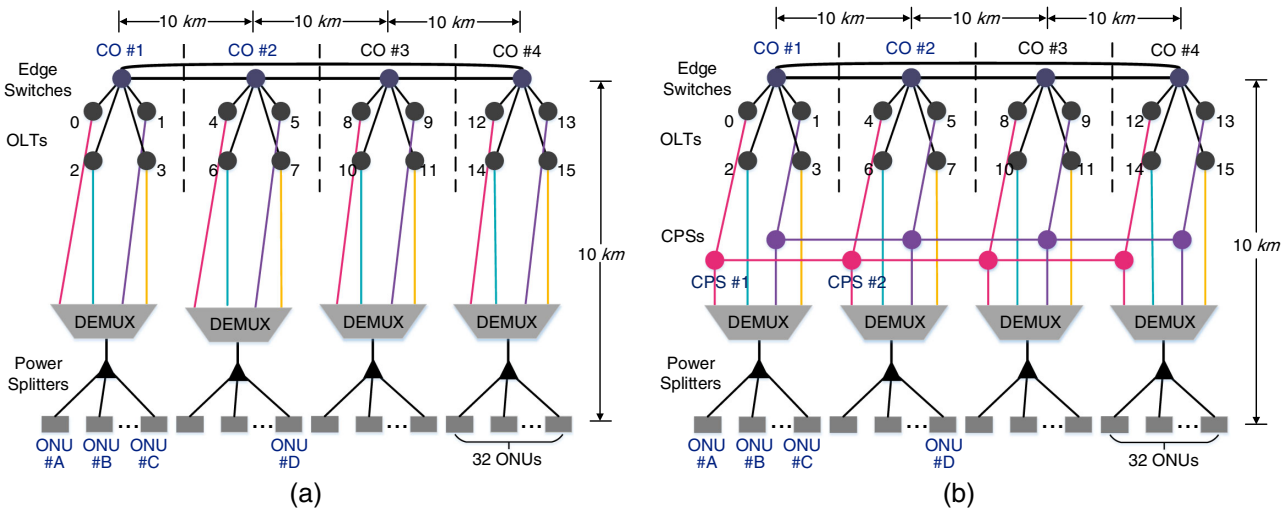

Fig. 10. Simulation topology. (a) TWDM-PON. (b) TS-TWDM-PON.

50 Mb/s and 100 Mb/s, respectively. In addition to considering pure TWDM-PON and TS-TWDM-PON scenarios, we also consider a "hybrid" scenario, with the coexistence of TWDM-PON and TS-TWDM-PON-II. In our simulation of the hybrid scenario, half of the wavelength resources are assigned to the TWDM-PON, while the remaining are assigned to the TS-TWDM-PON-II (indicated by TWDM-PON: TS-TWDM-PON-II = 1:1 in the figure). For a given traffic request, we first attempt to satisfy the request by using the WTSA algorithm to assign a wavelength resource belonging to TS-TWDM-PON; if that is unsuccessful, we use the First Fit algorithm to try to find an available wavelength among the TWDM-PON resources. When all wavelengths (including those of TS-TWDM-PON and TWDM-PON) cannot support the traffic, this request is blocked.

Figure 11(a) shows the total delay performance of the TWDM-PON architecture and the proposed TS-TWDM-PON under uniform traffic load. We observe that while the total delay of all four architectures increases with increasing traffic load, the delay of the TS-TWDM-PON scenario is much lower than that of the TWDM-PON. With respect to the three delay components, propagation delay is a physical property and cannot be optimized, regardless of the architecture applied (assuming the fiber distances involved are the same). The buffer delay difference between TS-TWDM-PON and TWDM-PON is relatively small with the value about 0.1 ms, which is discussed in detail in the next subsection. It is the processing delay that dominates the total delay, especially under high traffic load. However, while the processing delay is inevitable in a TWDM-PON because of the need for electronic switching for some of the traffic, it is avoidable in our approach because the inter-CO data flows are all-optically switched at the CPSs.

The total delay for TS-TWDM-PON-I and TS-TWDM-PON-II is the same, because the bandwidth granularity of the OTSS has virtually no impact on delay performance (the small difference in network throughput in the two scenarios was negligible, as will be shown later). Considering the hybrid architecture, we observe that when the traffic load is light, e.g., less than 30 Erlang, the total delay is very low and almost the same as that of TS-TWDM-PON. This is because, in the hybrid architecture, the requests are assigned to the TS-TWDM-PON-II wavelength resources if possible. If a request cannot be accommodated by TS-TWDM-PON-II, it uses the

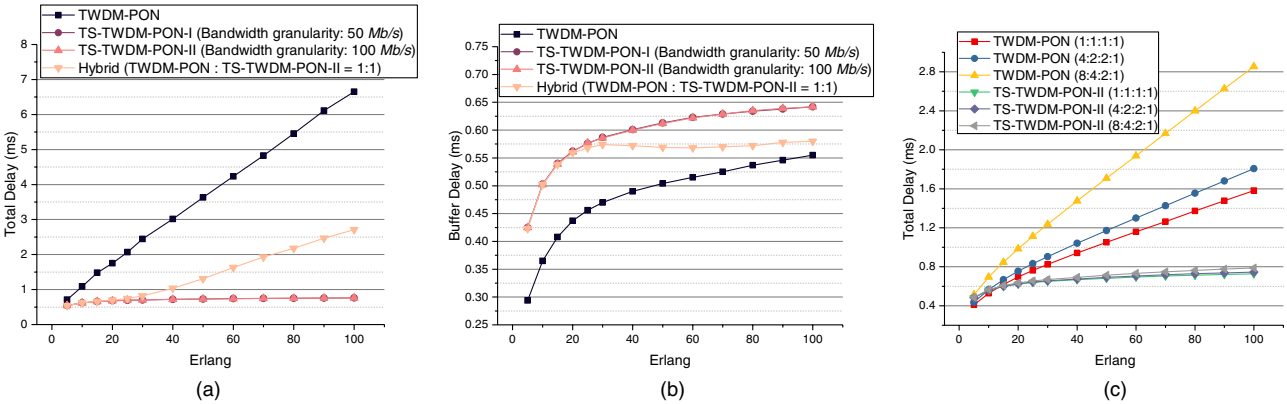

Fig. 11. (a) Total delay performances under uniform traffic load. (b) Buffer delay under different traffic loads. (c) Total delay performances under non-uniform traffic load.

wavelength resources assigned to the TWDM-PON, which significantly increases the latency.

C. Buffer Delay Under Different Traffic Loads

Next we investigate the buffer delay performance. To calculate the buffer delay, we need to know the exact arrival time of each Ethernet packet, the size of which ranges from 64 bytes to 1518 bytes. We apply the traffic model used in [39]. Self-similar traffic is generated by aggregating multiple sub-streams, each consisting of alternating Pareto-distributed ON/OFF periods. The continuous Ethernet packets are generated according to the bandwidth of the request. It is worth noting that the procedures used for time slot assignment in the TWDM-PON and TS-TWDM-PON are slightly different. In TS-TWDM-PON, a service request occupies contiguous optical time slots in the OTSS frames. The request transmits the data in the fixed positions of OTSS frames. This is because the CPSs operate periodically at the switching point after the start and end positions of the contiguous optical time slots are determined. In the TWDM-PON, the positions of the assigned time slots in the different frames need not be the same. The OLT can assign variable time slots in every cycle.

Figure 11(b) shows that the buffer delay increases with increasing traffic load. When the traffic load is light, the packet does not need to wait for a long time until it is transmitted. We notice that the buffer delay of TS-TWDM-PON is greater than that of the TWDM-PON by approximately 0.1 ms. When a packet arrives at the ONU, it cannot be transmitted immediately and must be stored in the buffer until the assigned time slots begin. The best case is when the packet arrival and the start of the assigned time slots coincide and the buffer delay is essentially zero. The worst case is when the packet arrives, the assigned time slots are just finished, and the packet must wait for a whole frame time. TWDM-PON can alleviate the waiting time when too many packets arrive in a very short time by using a variable time slot assignment.

Regarding the hybrid architecture, the buffer delay is almost the same as that of TS-TWDM-PON when the traffic load is light. However, when the traffic load is larger than 30 Erlang, the buffer delay of the hybrid architecture decreases slightly at first, and then increases slightly. There are two opposite influencing factors. On one hand, with the increment of the traffic load, more requests are accommodated by the TWDM-PON, and the average buffer delay is reduced. On the other hand, buffer delay increases with increasing traffic load. In this case, the average buffer delay of the hybrid architecture is increased.

D. Network Delay Performances Under Non-Uniform Traffic Load

As discussed in the Introduction, a user cannot always receive low-latency service because of the limited storage capacity of a single MDC. To simulate this realistic situation, we also analyze the network delay performance under a non-uniform traffic load. Unlike the previous simulation where the ONUs randomly selected one of the four COs as the destination, in this situation, an ONU selects the closest CO with sufficient resources in priority. Although the resources in each CO are equal, the number of traffic requests generated is different in each PON system. In the simulation, we designed three types of traffic loads for the four COs. The first was 1:1:1:1, which means that the proportion of the overall traffic load for all four COs is the same. The other two types were 4:2:2:1 and 8:4:2:1. Obviously, the third traffic load is the most unbalanced. For the evaluation, we used the same simulation topology depicted in Fig. 10.

We selected TWDM-PON and TS-TWDM-PON-II to compare the network delay performance under an unbalanced traffic load. As shown in Fig. 11(c), the more unbalanced the traffic load, the higher the total delay of TWDM-PON. The unbalanced traffic load leads to some COs being heavy loaded, while the others are relatively idle. Under such circumstances, a request is more likely to be served by remote MDCs, leading to higher propagation delay and excessive processing time. We notice that the total delay for TWDM-PON is slightly higher than for TS-TWDM-PON when the traffic load is light, i.e., less than 10 Erlang. It is because most of the requests are accommodated by the closest CO under light traffic load, regardless of the unbalanced traffic load distribution. It is the buffer delay that dominates the total delay; no processing delay is introduced under light traffic load conditions. As mentioned earlier, the buffer delay of TS-TWDM-PON is greater than that of the TWDM-PON by approximately 0.1 ms, leading to the small total delay difference between the two architectures under light traffic load. Furthermore, there is virtually no difference in the total delay of TS-TWDM-PON with any of three types of traffic load. This is because only propagation delay increases with the traffic load increment; excessive processing time is avoided; and the buffer delay is the same, as was mentioned earlier. Compared with the processing time, propagation delay is much lower, which is the reason that the differences are negligible.

E. Network Throughput Analysis

Figures 12(a) and 12(b) show the average network throughput every 100 s when the traffic load is 50 Erlang and 100 Erlang, respectively. The overall simulation time is 1000 s. We observe that the network throughput performance of the TWDM-PON architecture is higher than that of the other three approaches, while TS-TWDM-PON-II whose bandwidth granularity is 100 Mb/s exhibits the lowest network throughput. This is because OTSS requires contiguous optical time slots at different links; thus, some requests may not be accommodated despite other optical time slots being idle, which reduces the network throughput.

The network throughput of TS-TWDM-PON-I is slightly higher than that of TS-TWDM-PON-II because the bandwidth granularity of the former is finer, which results in higher probability of accommodating a request, resulting in higher network throughput. Let us consider an example where a traffic request needs 550 Mb/s bandwidth and we need to calculate the number of required time slots. The bandwidth granularity of TS-TWDM-PON-I is 50 Mb/s (length of a minimum optical time slice is 5 μs), and 11 contiguous minimum time slices are occupied (55 μs in total). Whereas, in TS-TWDM-PON-II with a bandwidth granularity of 100 Mb/s (length of a

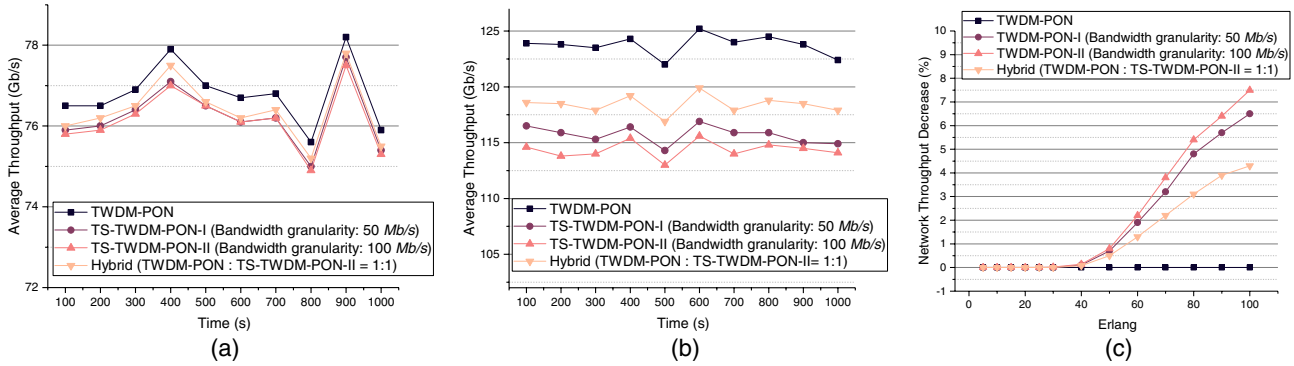

Fig. 12. (a) Average network throughput under 50 Erlang. (b) Average network throughput under 100 Erlang. (c) Network throughput decrease compared with TWDM-PON under different traffic loads.

minimum optical time slice is 10 μs), six continuous minimum optical time slices are occupied (60 μs in total). Thus, we see that TS-TWDM-PON-II may occupy more resources than TS-TWDM-PON-I while serving the same request.

Figure 12(c) shows the decrease in network throughput of TS-TWDM-PON when compared with TWDM-PON under different traffic loads. When traffic load is lower than 40 Erlang, the network throughput difference is almost zero. The maximum difference is around 7.5% at 100 Erlang, which is an extremely high traffic load condition. At 80 Erlang (also a relatively high traffic load), the loss is around 5%. Therefore, the network throughput decline for TS-TWDM-PON compared with TWDM-PON is minimal and acceptable.

6. EXPERIMENTAL VALIDATION

We also conducted a prototype experiment to validate the feasibility of the proposed network architecture. Figure 13 demonstrates the experimental setup. Four data flows, marked as #A, #B, #C, and #D, of lengths of 10 μs, 10 μs, 1 μs, and 25 μs, respectively, were generated. Note that these four ONUs and the destinations of the four data flows, namely, MDC #1 and MDC #2, are as depicted in Fig. 10. Using the topology in Fig. 10, we can have a better understanding of the whole experimental setup. ONUs #A, #B, and #C are from the same PON system, while ONU #D belongs to another PON. Data flows #A, #B, and #C can reach MDC #1 without passing through an edge switch, while data flow #D can reach MDC #2 in a similar manner. In our setting, the destination for data flows #A and #C is MDC #1, and that for data flows #B and #D is MDC #2. Note that the ONUs and CPSs are time-synchronized to ensure the correct switching operation of switches. Two CPSs are used in our experiment, which are realized by lead-lanthanum-zirconate-titanate (PLZT) switches.

The following is a detailed discussion of the data flow transmission. To emulate the PON architecture, data flows #A, #B, and #C go through a wavelength-division multiplex (WDM) multiplexer (MUX) and demultiplexer (DEMUX) after they are generated. Data flow #A is directly sent to MDC #1, while #B and #C are sent to CPS #1. After passing through CPS #1, data flow #B goes through CPS #2, along with data flow #D. The digital storage oscilloscope (DSO) shows the received data from different ports.

Figure 14 presents the experimental results. The lengths of the OTSS frame and the guard interval are set to 125 μs and

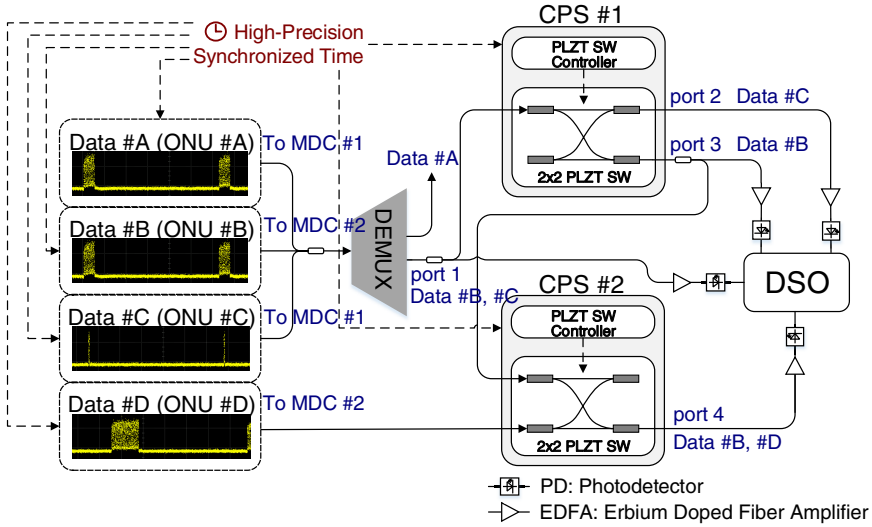

Fig. 13. Experiment setup.

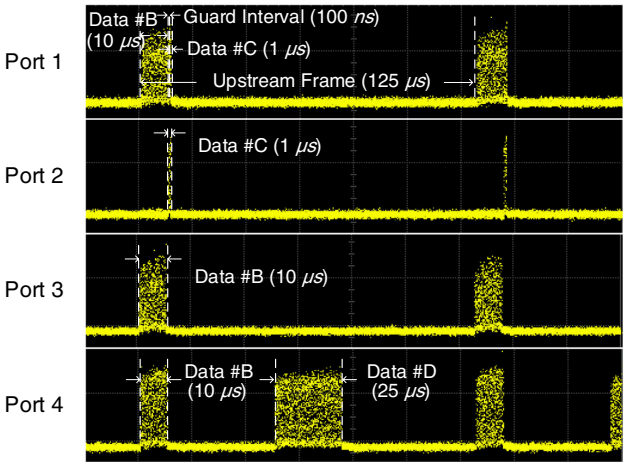

Fig. 14. Experimental results.

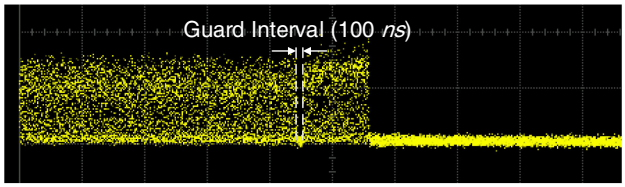

Fig. 15. Guard interval with 100 ns.

100 ns, respectively. The guard interval corresponds to the 1588v2 protocol with sub-microsecond-level time synchronization accuracy. A clearer figure of the guard interval with 100 ns is shown in Fig. 15. As shown in Fig. 14, data flows are transmitted on repetitive OTSS frames in the time domain. After passing through a WDM DEMUX at port 1, data flows #B and #C are separated at CPS #1 and switched to port 3 and port 2, which represent the data received at MDC #2 and MDC #1, respectively. In port 4, data flows #B and #D are aggregated correctly after going through CPS #2. Thus, we achieved the separation of the optical time slices at CPS #1 and the combination at CPS #2. This experiment successfully demonstrates that by applying OTSS, the data flows from different PONs can be accurately transmitted without collision with the help of time synchronization technology, which validates the feasibility of the network architecture.

7. CONCLUSION

Edge computing has been regarded as a promising technology that can provide extremely low-latency connections for delay-sensitive applications. However, users cannot always be served by the closest MDC, due to the limited computing and storage resources available at a single MDC. To solve this problem, we proposed a flexible low-latency metro-access converged network architecture based on OTSS. By leveraging the transparent (i.e., all-optical) connections of OTSS, requests can be processed in other MDCs or cloud data centers without requiring extra processing time in electronic switches. The simulation results demonstrated that our proposed architecture could achieve much lower-latency connections than TWDM-PON under a range of conditions. Furthermore, we investigated the network throughput performances for different traffic loads, and the results showed that the network throughput of TS-TWDM-PON was just slightly lower than that of the TWDM-PON, when the traffic load was heavy, which is acceptable. We further conducted experiments to successfully validate the feasibility of the proposed architecture.

Funding. National Key R&D Program of China (2018YFB1801702); National Natural Science Foundation of China (61621064, 61871448).

Acknowledgment. Parts of this work were presented at the Optical Fiber Communication Conference (OFC), San Diego, USA, March 2018.

REFERENCES

1. A. Pizzinat, P. Chanclou, F. Saliou, and T. Diallo, "Things you should know about fronthaul," J. Lightwave Technol. **33**, 1077–1083 (2015).
2. 3GPP, "Release-16," https://www.3gpp.org/release-16.
3. M. Ruffini, "Multidimensional convergence in future 5G networks," J. Lightwave Technol. **35**, 535–549 (2016).
4. V. Bobrovs, S. Spolitis, and G. Ivanovs, "Latency causes and reduction in optical metro networks," Proc. SPIE **9008**, 90080C (2014).
5. K. Chen, A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen, and Y. Chen, "OSA: an optical switching architecture for data center networks with unprecedented flexibility," IEEE/ACM Trans. Netw. **22**, 498–511 (2013).
6. M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Toward low-latency and ultra-reliable virtual reality," IEEE Netw. **32**, 78–84 (2018).
7. Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," J. Internet Serv. Appl. **1**, 7–18 (2010).
8. P. Mach and Z. Becvar, "Mobile edge computing: a survey on architecture and computation offloading," IEEE Commun. Surv. Tutorials **19**, 1628–1656 (2017).
9. Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—a key technology towards 5G," ETSI white paper 11 (2015).
10. F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing* (ACM, 2012), pp. 13–16.
11. W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: vision and challenges," IEEE Internet Things J. **3**, 637–646 (2016).
12. E. Ahmed, A. Ahmed, I. Yaqoob, J. Shuja, A. Gani, M. Imran, and M. Shoaib, "Bringing computation closer toward the user network: is edge computing the solution?" IEEE Commun. Mag. **55**(11), 138–144 (2017).
13. Z. Zhong, N. Hua, M. Tornatore, Y. Li, H. Liu, C. Ma, Y. Li, X. Zheng, and B. Mukherjee, "Energy efficiency and blocking reduction for tidal traffic via stateful grooming in IP-over-optical networks," J. Opt. Commun. Netw. **8**, 175–189 (2016).
14. M. Aazam and E. N. Huh, "Dynamic resource provisioning through fog micro datacenter," in *IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)* (IEEE, 2015), pp. 105–110.
15. N. Hua and X. Zheng, "Optical time slice switching (OTSS): an all-optical sub-wavelength solution based on time synchronization," in *Asia Communications and Photonics Conference (ACP)* (2013), paper AW3H-3.
16. D. A. Khotimsky, "NG-PON2 transmission convergence layer: a tutorial," J. Lightwave Technol. **34**, 1424–1432 (2016).
17. Y. Luo, X. Zhou, F. Effenberger, X. Yan, G. Peng, Y. Qian, and Y. Ma, "Time-and wavelength-division multiplexed passive optical network

(TWDM-PON) for next-generation PON stage 2 (NG-PON2)," J. Lightwave Technol. **31**, 587–593 (2012).
18. N. Cheng, J. Gao, C. Xu, B. Gao, X. Wu, D. Liu, L. Wang, X. Zhou, H. Lin, and F. Effenberger, "World's first demonstration of pluggable optical transceiver modules for flexible TWDM PONs," in *39th European Conference and Exhibition on Optical Communication (ECOC)* (IET, 2013), pp. 1–3.
19. Y. Luo, M. Sui, and F. Effenberger, "Wavelength management in time and wavelength division multiplexed passive optical networks (TWDM-PONs)," in *IEEE Global Communications Conference (GLOBECOM)* (IEEE, 2012), pp. 2971–2976.
20. Optoplex, "10Gbps tunable ROSA for NG-PON2," https://www.optoplex.com/10G_Tunable_ROSA.htm.
21. Hisense, "NG-PON2 transceiver," http://www.hisensebroadband.com/html/Solutions/Access/NGPON_2/index.html.
22. P. Vetter, "Programmable access and edge cloud architecture," in *Optical Fiber Communication Conference (OFC)* (2017), paper M3I-4.
23. B. Yang, Z. Zhang, K. Zhang, and W. Hu, "Integration of micro data center with optical line terminal in passive optical network," in *21st OptoElectronics and Communications Conference (OECC)* (IEEE, 2016), pp. 1–3.
24. S. Yao, B. Mukherjee, and S. Dixit, "Advances in photonic packet switching: an overview," IEEE Commun. Mag. **38**(2), 84–94 (2000).
25. C. Qiao and M. Yoo, "Optical burst switching (OBS)—a new paradigm for an optical Internet," J. High Speed Netw. **8**, 69–84 (1999).
26. L. Xu, H. G. Perros, and G. Rouskas, "Techniques for optical packet switching and optical burst switching," IEEE Commun. Mag. **39**(1), 136–142 (2001).
27. Y. Jia, N. Hua, Y. Li, and X. Zheng, "Applying multi-controller collaboration in fine-grained all-optical intra-datacenter networks," J. Opt. Commun. Netw. **10**, B37–B48 (2018).
28. Y. Li, N. Hua, and X. Zheng, "Fine-grained all-optical switching based on optical time slice switching for hybrid packet-OCS intra-data center networks," in *Optical Fiber Communication Conference (OFC)* (Optical Society of America, 2016), paper W3J-5.
29. "Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications," IEEE Std 802.3ah (2000).
30. Y. Luo, F. J. Effenberger, and N. Ansari, "Time synchronization over Ethernet passive optical networks," IEEE Commun. Mag. **50**(10), 136–142 (2012).
31. L. Han, H. Li, L. Wang, N. Hua, C. Hu, J. Wang, S. Liu, L. He, Z. Chen, and Y. Xu, "First national high-precision time synchronization network with sub-microsecond accuracy over commercial optical networks for wireless applications," in *Asia Communications and Photonics Conference (ACP)* (2014), paper AF4B-6.
32. "IEEE standard for a precision clock synchronization protocol for networked measurement and control systems," IEEE Std 1588-2008 (2008).
33. M. Lévesque and D. Tipper, "A survey of clock synchronization over packet-switched networks," IEEE Commun. Surv. Tutorials **18**, 2926–2947 (2016).
34. R. Luo, N. Hua, X. Zheng, and B. Zhou, "High-reliability sub-nanosecond network time synchronization method enabled by double-frequency distributed time synchronization," J. Opt. Commun. Netw. **11**, A40–A51 (2019).
35. S. Lee, "An enhanced IEEE 1588 time synchronization algorithm for asymmetric communication link using block burst transmission," IEEE Commun. Lett. **12**, 687–689 (2008).
36. B. Ning, P. Du, D. Hou, and J. Zhao, "Phase fluctuation compensation for long-term transfer of stable radio frequency over fiber link," Opt. Express **20**, 28447–28454 (2012).
37. L. Yu, R. Wang, L. Lu, Y. Zhu, C. Wu, B. Zhang, and P. Wang, "Stable radio frequency dissemination by simple hybrid frequency modulation scheme," Opt. Lett. **39**, 5255–5258 (2014).
38. S. B. Yoo, "Optical packet and burst switching technologies for the future photonic Internet," J. Lightwave Technol. **24**, 4468–4492 (2006).
39. L. Wang, X. Wang, B. Mukherjee, H. S. Chung, H. H. Lee, and S. Park, "On the performance of hybrid-PON scheduling strategies for NG-EPON," in *International Conference on Optical Network Design and Modeling (ONDM)* (IEEE, 2016), pp. 1–5.