Energy Efficiency of an Integrated Intra-Data-Center and Core Network With Edge Caching

Matteo Fiorani, Slavisa Aleksic, Paolo Monti, Jiajia Chen, Maurizio Casoni, and Lena Wosinska

***Abstract*—The expected growth of traffic demand may lead to a dramatic increase in the network energy consumption, which needs to be handled in order to guarantee scalability and sustainability of the infrastructure. There are many efforts to improve energy efficiency in communication networks, ranging from the component technology to the architectural and service-level approaches. Because data centers and content delivery networks are responsible for the majority of the energy consumption in the information and communication technology sector, in this paper we address network energy efficiency at the architectural and service levels and propose a unified network architecture that provides both intra-data-center and inter-data-center connectivity together with interconnection toward legacy IP networks. The architecture is well suited for the *carrier cloud* model, where both data-center and telecom infrastructure are owned and operated by the same entity. It is based on the hybrid optical switching (HOS) concept for achieving high network performance and energy efficiency. Therefore, we refer to it as an integrated HOS network. The main advantage of the integration of core and intra-data-center networks comes from the possibility to avoid the energy-inefficient electronic interfaces between data centers and telecom networks. Our results have verified that the integrated HOS network introduces a higher number of benefits in terms of energy efficiency and network delays compared to the conventional nonintegrated solution. At the service level, recent studies demonstrated that the use of distributed video cache servers can be beneficial in reducing energy consumption of intra-data-center and core networks. However, these studies only take into consideration conventional network solutions based on IP electronic switching, which are characterized by relatively high energy consumption. When a more energy-efficient switching technology, such as HOS, is employed, the advantage of using distributed video cache servers becomes less obvious. In this paper we evaluate the impact of video servers employed at the edge nodes of the integrated HOS network to understand whether edge caching could have any benefit for carrier cloud operators utilizing a HOS network architecture. We have demonstrated that if the distributed video cache servers are not properly dimensioned they may have a negative impact on the benefit obtained by the integrated HOS network.**

***Index Terms*—Backbone networks; Edge caching; Energy consumption; Hybrid optical switching; Intra-data-center networks; Performance analysis.**

Manuscript received September 27, 2013; revised February 15, 2014; accepted February 17, 2014; published March 31, 2014 (Doc. ID 198202).

M. Fiorani (e-mail: matteo.fiorani@unimore.it) and M. Casoni are with the Department of Engineering "Enzo Ferrari," University of Modena and Reggio Emilia, Modena, Italy.

S. Aleksic is with the Institute of Telecommunications, Vienna University of Technology, Vienna, Austria.

P. Monti, J. Chen, and L. Wosinska are with the KTH Royal Institute of Technology, Kista, Sweden.

http://dx.doi.org/10.1364/JOCN.6.000421

I. Introduction

Although information and communication technology (ICT) can play a fundamental role in enabling a low-carbon economy, the energy and carbon impact of the ICT sector itself is already significant, and it is expected to grow rapidly with the proliferation of the number of connected devices and with the emergence of new services. The energy consumption of the ICT sector can be divided into (i) energy consumed by the user devices, (ii) energy consumed by the telecommunication network infrastructure, and (iii) energy consumed by the data centers. While end user devices are the major contributors, the sum of the energy consumed by the telecommunication networks and data centers amounts to 51% [1] of the total ICT consumption. With the expected growth in the Internet and data center traffic [2,3] the energy consumption of telecommunication networks and data centers is destined to drastically increase if the network energy efficiency is not improved. In addition to low-power device technologies, this problem can be addressed on architectural and service levels.

Taking into consideration the architectural level, we observe that generally telecommunication networks can be divided into three areas: access, metro, and core. Several research papers address the energy consumption of the different network areas [4,5]. It was shown that although access networks are currently the major contributor, the energy consumption of core networks is expected to grow rapidly to be able to support very high capacities in the range of several hundreds of terabits per second or even petabits per second per node [2]. As for data centers, their energy consumption is divided into energy consumed by the information technology (IT) equipment, energy consumed by the cooling system, and energy consumed by the power supply chain. According to the latest specifications, data centers are designed in such a way that the ICT equipment consumes nearly all the energy within

1943-0620/14/040421-12$15.00/0

the data center. This also means that in modern data centers major energy savings can be achieved by reducing the power consumption of the IT equipment. According to [6], the intra-data-center network, which handles the traffic inside a data center as well as that destined to the external networks, currently represents 23% of the IT equipment energy consumption. This number is expected to grow in the future due to the forecasted increase in the data center traffic [3]. It is, therefore, of the utmost importance to define new energy-efficient intra-data-center network technologies. In [5,7] it has been shown that the switching infrastructure consumes the major part of the energy in core and intra-data-center networks, and it was pointed out that future research needs to focus on improving the energy efficiency of switching devices. Today, core and intra-data-center networks are based on electronic switching; that is, data transmission is performed in the optical domain, whereas switching and control is done by electronic equipment. Consequently, electrical-to-optical (E/O) and optical-to-electrical (O/E) conversions are performed at each hop, which leads to high energy consumption. To solve this problem, several optical switching solutions have been proposed in core [8,9] and intra-data-center [10,11] networks. In particular, [9,11] proposed two architectures based on hybrid optical switching (HOS) for achieving high performance and energy efficiency in core and intra-data-center networks, respectively. The term *hybrid* is used to describe the coexistence of different optical switching paradigms, namely packet, burst, and circuit switching.

Meanwhile, the latest Cisco Visual Networking Index forecast [2] reports that consumer Internet video traffic will increase from 57% to 69% of total Internet traffic in the period between 2012 and 2017. As a consequence, energy-efficient video distribution systems are an important tool to maintain a sustainable Internet growth. Video content can be either stored and distributed from a few centralized servers located in large data centers (referred to as the centralized approach) or the most popular video contents can be replicated in cache servers located close to the end users (referred to as the distributed approach). From the energy consumption perspective it is not obvious which approach (i.e., centralized or distributed) is most beneficial. In fact, storing content only in a centralized server decreases the energy consumption for storage while increasing transport energy requirements. On the other hand, replicating some content closer to the users in distributed cache servers decreases transport energy while increasing the storage energy requirements. A few recent studies [12–15] address this trade-off and conclude that the highest energy efficiency is achieved by storing popular content in cache servers close to the end users.

Recently, communication service providers are looking for cloud solutions to reduce costs and create a new level of efficiency. In this context, one of the most promising solutions is the carrier cloud model, where both data centers and the core network are owned by the same entity and the resources are virtualized and shared by multiple tenants. Several large telecom operators are considering a move to this novel business model [16,17]. Carrier clouds could overcome several problems that occur in the existing cloud solutions, such as unpredictable and nondeterministic network performance and insufficient availability and security, which severely complicate or even preclude carrier-grade service level agreements. In order to increase both the adaptability to different traffic types and the energy efficiency at the architectural level, this paper proposes a unified network architecture for carrier cloud operators. The architecture is based on HOS and provides both intra-data-center and inter-data-center connectivity as well as interconnection capabilities toward legacy IP networks. This architecture is referred to as an integrated HOS network, in which the traffic is carried in the optical domain along the entire path from an aggregation switch inside a data center up to another aggregation switch (in the same or a different data center) or to an edge node serving as an interface to legacy IP networks. In our study, we analyze the structure of such an integrated architecture and evaluate the benefits compared to a nonintegrated HOS architecture as well as a conventional IP network based on electronic switching.

Regarding the service level, we observe that the studies in [12–15] take into consideration only traditional core and intra-data-center networks based on electronic switches. Since these networks are characterized by a low energy efficiency, the reduction in the transport energy introduced by the distributed storing approach generally overcomes by far the energy consumption of the cache servers. If we consider instead a carrier cloud operator that relies on the integrated HOS network, which is able to achieve high energy efficiency, the advantage of the distributed approach on energy consumption might become less obvious. For this reason, in this paper we deploy distributed cache servers at the edge nodes of the integrated HOS network and evaluate their impact on the network performance and energy consumption. To the best of our knowledge, the performance of edge caching in combination with an energy-efficient network concept based on optical switching has not been evaluated so far.

To summarize, the contribution of this paper is twofold, namely, (i) we propose and evaluate an integrated intra-data-center and core network architecture based on the HOS concept for carrier cloud operators along with a study of its benefits and (ii) we assess the impact of distributed video cache servers on the proposed integrated HOS network architecture.

The remainder of the paper is organized as follows. In Section II we describe the proposed integrated core and intra-data-center HOS network. Section III introduces the approach used to model the video cache servers. In Section IV the reference network used for the simulations and the energy consumption model are described. Section V presents the simulation results, while Section VI contains some concluding remarks.

II. Integrated Intra-Data-Center and Core Network

Figure 1 shows a high-level representation of the proposed integrated core and intra-data-center network based

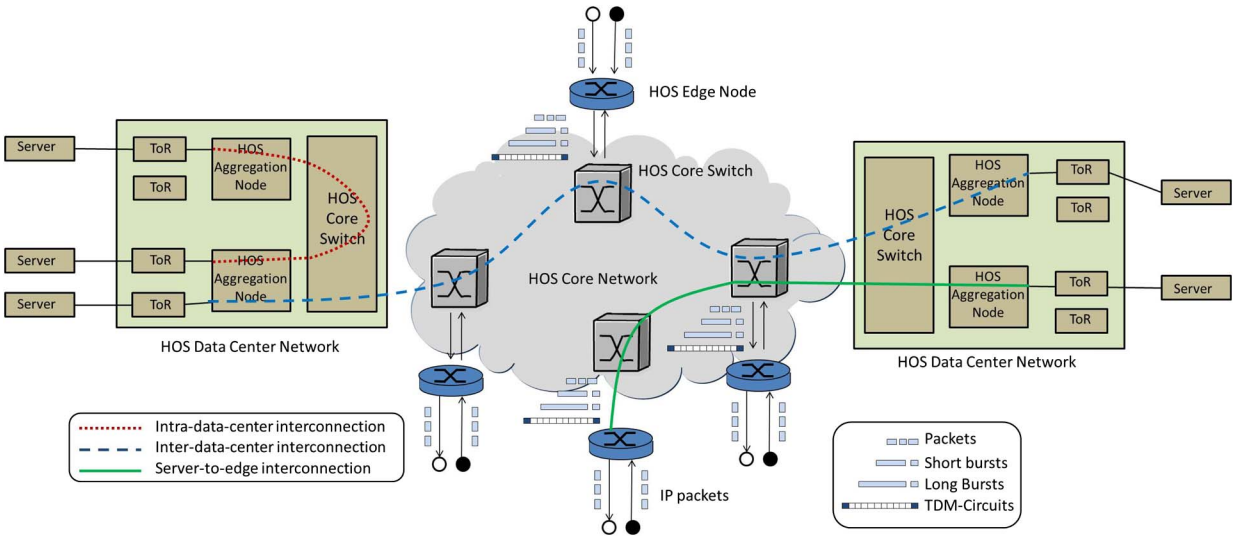

Fig. 1. Representation of the proposed integrated intra-data-center and core network based on hybrid optical switching.

on HOS. The integrated network provides three different types of interconnections using a unified all-optical infrastructure and a common control plane. The first type of interconnection is between servers inside the same data center, referred to as an *intra-data-center interconnection*. In Fig. 1 it is represented by a red dotted line to highlight the path over which data are sent using the HOS paradigm. The second type of interconnection is between servers located in different data centers. We refer to it as an *inter-data-center interconnection*, and we use a blue dashed line in Fig. 1 to indicate the path performed in the HOS domain. The third type of interconnection is between servers inside a data center and HOS edge nodes; that is, it provides the *server-to-edge interconnections*. An example is indicated in Fig. 1 by a green solid line for the HOS path.

It should be noted that in the proposed integrated network, core and data centers employ the same unified control plane. A first attempt of an integrated control plane for intra-data-center and core networks for a carrier cloud has been recently proposed in [17]. The authors create a proof of concept for an integrated control plane based on the software-defined network mechanism. However, the proposed solution is still based on traditional electronic interfaces between data centers and core networks, and hence it is not optimized from the energy-efficiency point of view. In order to minimize energy consumption, in this paper we propose a novel control plane for intra-data-center and core networks based on the HOS network model described in [9,11]. It consists of two layers, the generalized multiprotocol label switching (GMPLS) control layer and the HOS forwarding layer. The GMPLS control layer is in charge of configuring and managing the network virtual topology. It consists of three building blocks: routing, signaling, and link management. The HOS forwarding layer performs data aggregation, data scheduling, and resource reservation. It supports three different optical transport mechanisms, namely circuits, bursts, and packets. The HOS forwarding layer has the unique feature of employing a common control packet for managing all three switching paradigms, enabling circuits, bursts, and packets to dynamically share the optical resources. The use of optical bursts in combination with packets and circuits allows the dynamic implementation of different service classes, leading to an efficient quality-of-service differentiation.

A. HOS Core Network

The HOS core network provides connectivity among different data centers as well as between data centers and legacy IP networks. As shown in Fig. 1, each node in the HOS core network includes a HOS core switch. If the node is located at the edge of the HOS core network, it is equipped with an electronic switch for interdomain connectivity.

An electronic switch in the HOS edge node ensures interoperability between the core network and the legacy IP networks. In the direction toward the HOS core network, the HOS edge node performs traffic classification and traffic aggregation. In other words, each incoming IP packet is classified based on the value of the differentiated service code point field in the IP header and mapped over the best-suited optical transport mechanism, as described in [9]. In the direction toward the legacy IP networks, the HOS edge node extracts IP packets and performs IP routing. The HOS edge node is divided into two logical building blocks, one of which consists of an electronic switch to perform IP routing and the second one of which includes all the electronic components required to (i) perform traffic aggregation and classification in the direction toward the HOS core network and (ii) perform IP packet extraction in the direction toward the legacy IP networks. For simplicity, we will refer to this block as the traffic aggregation block.

High-capacity optical switches provide connectivity inside the core network. A HOS core switch can be logically divided into two building blocks, namely the electronic control logic and the optical switching fabric. The electronic control logic consists of three electronic blocks for

implementing the GMPLS control layer, the HOS forwarding layer, and the switch control unit. The optical switching fabric is composed of two large optical switches. A fast optical switch, based on semiconductor optical amplifiers (SOAs), takes care of the transmission of packets and short bursts. A slow optical switch, based on microelectromechanical systems (MEMS), handles the transmission of circuits and long bursts. In the optical switching fabric block we also include the following active optical components: optical amplifiers (OAs), tunable wavelength converters (TWCs), and control information extraction/reinsertion (CIE/R) blocks in order to compensate for signal losses in components, reduce blocking probability, and encode the control information together with the data payload on the same optical carrier, respectively.

For a detailed description of the HOS core network we refer to [9].

B. HOS Intra-Data-Center Network

The HOS intra-data-center network provides connectivity among the servers inside a data center and connects the data centers to the HOS core network. It is organized in a three-tier fat-tree topology. The first tier consists of electronic top-of-rack (ToR) switches. In a conventional high-end data center, servers are organized in racks, with each rack hosting typically 48 blade servers. The ToR switches interconnect the servers inside a rack and connect the racks to the second tier of the intra-data-center network, which is composed of the HOS aggregation nodes. The HOS aggregation nodes perform the same functions inside a data center as HOS edge nodes in the HOS core network. In particular, in the direction toward the network core, the HOS aggregation nodes perform traffic classification and traffic aggregation, while in the direction toward the data center servers, the HOS aggregation nodes extract the IP packets and perform IP routing. The HOS aggregation nodes consist of the same logical building blocks as the HOS edge nodes. The main difference between HOS edge and HOS aggregation nodes is that the HOS edge nodes could also include the video cache servers, which will be further elaborated in Section III. The third tier of the intra-data-center network is represented by a single large HOS core node. This node has exactly the same architecture as the HOS core switch used in the core network. For more details we refer to [11].

III. Edge Cache

To evaluate the impact of distributed video cache servers on the proposed integrated intra-data-center and core HOS network, we extend the HOS edge node architecture described in Section II in order to include the video cache servers. The extended architecture of HOS edge nodes with cache servers is shown in Fig. 2. It can be logically divided into three building blocks. Two of them have already been mentioned in Section II, namely the electronic switch block and the traffic aggregation block. The former one includes the switch, the GMPLS module, and the input electronic line cards, while the latter block comprises the classifier, the conditioner, the assembler, the resource allocator, and the packet extractor. The last block, which represents an extension to the architecture previously presented in [9,11], is related to the caching operations and consists of the content tracker, the ToR switch, and the video cache servers. The content tracker interacts with the HOS control plane in order to keep track of all the video content inside the cache servers, process the incoming video requests, and update the cache servers.

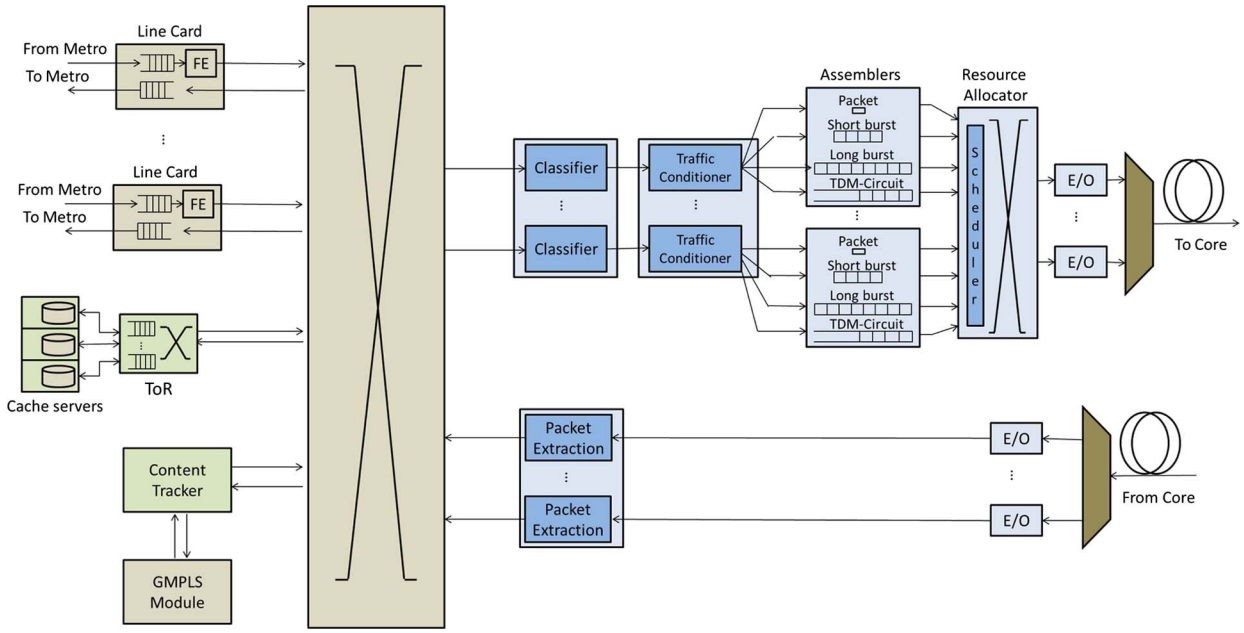

Fig. 2. Architecture of the HOS edge node with video cache servers.

As already mentioned in Section I, the impact of distributed cache servers on the network energy efficiency has already been addressed in previous studies [12–15], mainly focused on electronic switched networks. The rationale behind these works is that distributed cache servers reduce the traffic load, leading to a lower number of electronic switch ports used in the core and intra-data-center networks. However, the electronic switching devices commercially available today do not implement dynamic switching-off of the ports, and thus their energy consumption is almost independent of the traffic load. Techniques for dynamically switching off the line cards (LCs) have been proposed in [18,19], but their efficiency in real network scenarios has still to be proven. In fact, scheduling the switching-off of the LCs in a packet switching network is a very challenging task because of the stochastic nature of the traffic, and usually the interarrival time between two successive packets is very small. The novelty of our approach consists in applying the caching concept to a HOS network, where we assume that all the optical components (in the optical switching fabric of the HOS core nodes) are turned off when they are inactive. This is not as challenging as turning off electronic switch ports [9]. In fact, with two parallel optical switches, only one needs to be active to serve traffic from a particular port at a specified time. In addition, in a HOS network, circuits and bursts are scheduled *a priori*; thus the incoming traffic is more predictable than in a traditional packet switched network, that is, one where the traffic is processed on a packet-by-packet basis.

IV. Modeling Approach

In this section, we describe assumptions used to model and evaluate performance of the proposed integrated intra-data-center and core HOS network with edge caching. First, we present the power consumption model followed by the description of the reference network scenario, and finally we introduce the performance metrics.

A. Power Consumption Model

The total power consumption of the integrated core and intra-data-center HOS network is given by the sum of the power consumed by each node in the core network (P^i_{Node}) and the power consumed by data centers (P^j_{DC}):

$$P_{\text{Network}} = \sum_{i=1}^{N_{\text{Node}}} P^i_{\text{Node}} + \sum_{j=1}^{N_{\text{DC}}} P^j_{\text{DC}}, \tag{1}$$

where N_{Node} is the number of nodes in the core network and N_{DC} is the number of data centers. Each node in the HOS core network performs both edge and core functions. The power consumption of the ith node in the network is determined by

$$P^i_{\text{Node}} = P^i_{\text{Edge}} + P^i_{\text{Core}}, \tag{2}$$

where P^i_{Edge} is the power consumption of the ith HOS edge part and P^i_{Core} is the power consumption of the ith HOS core switch. The power consumption of the ith HOS edge part is given by the sum of the power consumption of its building blocks (Section II):

$$P^i_{\text{Edge}} = N_F^{\text{Edge},i} \cdot N_W \cdot (P_{\text{ES}} + P_A) + P^i_{\text{Cache}}, \tag{3}$$

where $N_F^{\text{Edge},i}$ is the total number of fibers connected to the HOS edge node i and N_W is the number of wavelengths per fiber, which are assumed to be the same for all nodes. In the formula, P_{ES} is the power consumption of the electronic switch block per port and P_A is the power consumption of the traffic aggregation block per port. The number of ports of the switch is given by the product of the number of wavelength channels per fiber and the number of fibers ($N_F^{\text{Edge},i} \cdot N_W$); that is, it represents the total number of wavelength channels at a HOS edge node. Finally, P^i_{Cache} is the power consumption of the cache block. The power consumption of the cache block of the ith HOS edge node is obtained through the following formula:

$$P^i_{\text{Cache}} = P_{\text{CT}} + P_{\text{ToR}} + N^i_{\text{CS}} \cdot P_{\text{CS}}, \tag{4}$$

where P_{CT} is the power consumption of the content tracker, P_{ToR} is the power consumption of the ToR switch, and P_{CS} is the power consumption of a cache server. Finally, N^i_{CS} represents the number of cache servers hosted in the ith HOS edge node. The cache servers are assumed to have a fixed storage capacity of 1 TByte. Also the power consumption of the ith HOS core switch is computed by summing up the power consumption of its building blocks, as defined by Eq. (5):

$$P^i_{\text{Core}} = P^i_{\text{ECL}} + P^i_{\text{OSF}}, \tag{5}$$

where P^i_{ECL} is the power consumption of the electronic control logic and P^i_{OSF} is the power consumption of the optical switching fabric of the ith HOS core switch. The power consumption of the control logic of the ith HOS core switch is given by Eq. (6):

$$P^i_{\text{ECL}} = N_F^{\text{Core},i} \cdot N_W \cdot P_{\text{GMPLS}} + P_{\text{HOS}} + P_{\text{SC}}, \tag{6}$$

where $N_F^{\text{Core},i}$ is the total number of fibers connected to the HOS core node i. In Eq. (6), P_{GMPLS} is the power consumption of the GMPLS block per port, P_{HOS} is the power consumption of the HOS forwarding layer, and P_{SC} is the power consumption of the switch control unit. The power consumption of the optical switching fabric of HOS core nodes depends on the traffic because we assume that optical switch ports can be turned off when they are inactive. To compute the power consumption of the optical switching fabric of the ith HOS core switch, we use Eq. (7):

$$P^i_{\text{OSF}} = N^{\text{active},i}_{\text{SOA}} \cdot P_{\text{SOA}} + N^{\text{active},i}_{\text{MEMS}} \cdot P_{\text{MEMS}} + N^{\text{active},i}_{\text{TWC}} \cdot P_{\text{TWC}} + N_F^{\text{core},i} \cdot (N_W \cdot P_{\text{CIE/R}} + 2 \cdot P_{\text{EDFA}}). \tag{7}$$

Here, $N^{\text{active},i}_{\text{SOA}}$, $N^{\text{active},i}_{\text{MEMS}}$, and $N^{\text{active},i}_{\text{TWC}}$ represent the number of active SOA-switch ports, MEMS-switch ports, and TWCs of the ith HOS core node, respectively. These values depend

on the traffic load and are computed through simulations. In Eq. (7), P_{SOA}, P_{MEMS}, and P_{TWC} are the power consumption of the SOA-switch per port, the MEMS-switch per port, and the TWC, respectively. Finally, $P_{\mathrm{CIE/R}}$ and P_{EDFA} are the power consumption of the CIE/R block and the OAs. When computing the power consumption of the HOS intra-data-center networks we exclude from our analysis the power consumed by servers and consider only the power consumed by the network equipment, that is, by the intra-data-center network. The power consumption of the jth intra-data-center network is computed using Eq. (8):

$$P^j_{\mathrm{DC}} = N^j_{\mathrm{ToR}} \cdot P_{\mathrm{ToR}} + N^j_{\mathrm{Aggr}} \cdot P_{\mathrm{Agrr}} + P^j_{\mathrm{Core}}, \tag{8}$$

where N^j_{ToR} and N^j_{Aggr} are the numbers of ToR switches and HOS aggregation switches in the jth data center, respectively. Here, P_{Aggr} represents the power consumption of an HOS aggregation switch and P^j_{Core} represents the power consumption of the HOS core switch inside the jth data center. We assume that each HOS aggregation switch is connected to the corresponding HOS core switch in the data center using one fiber and that the number of ToR switches connected to the corresponding aggregation node is equal to the number of wavelength channels per fiber (N_W). To calculate the power consumption of a HOS aggregation switch we use Eq. (9):

$$P_{\mathrm{Aggr}} = N_W \cdot (P_{\mathrm{ES}} + P_A). \tag{9}$$

Finally, the power consumptions of the HOS core switch inside the jth data center is computed using Eq. (5) and replacing the index i with the index j. The energy consumptions of all the considered network components are reported in Table I and have been obtained by collecting data from data sheets as well as from research papers [9,11].

TABLE I
POWER CONSUMPTION OF THE NETWORK COMPONENTS [9,11]

Components	Power [W]
Electronic switching block per port (P_{ES})	320
Traffic aggregation block per port (P_A)	159
Content tracker (P_{CT})	330
Top-of-rack switch (P_{ToR})	650
Cache server (P_{CS})	450
GMPLS control layer per port (P_{GMPLS})	6.75
HOS forwarding layer (P_{HOS})	570
Switch control unit (P_{SC})	300
SOA switch per port (P_{SOA})	20
MEMS switch per port (P_{MEMS})	0.1
Tunable WC (P_{TWC})	1.69
Control information E./R. ($P_{\mathrm{CIE/R}}$)	17
Optical amplifiers (P_{EDFA})	14

B. Reference Network Scenario

To assess the performance of the proposed integrated intra-data-center and core HOS network with edge caching, we developed a custom event-driven C++ simulator. In the following we report the main parameters that we used in our simulations and present the model that is applied to generate the network traffic.

We denote N_{Node} as the number of nodes in the network and N_{DC} as the number of nodes connected to the data center. We consider the Pan-European network [20] composed of 28 nodes (i.e., $N_{\mathrm{Node}} = 28$) and 41 links as the reference network topology. We assume that 25% of the network nodes are connected to a data center, that is, $N_{\mathrm{DC}} = 7$. In each simulation we randomly connect the data centers to different nodes of the network. We assume that all the data centers have the same size and are equipped with 76,800 servers organized in racks. In each rack, 48 servers are connected to a ToR switch using dedicated 1 Gbps links [21]. The number of ToR switches per data center is given by the ratio between the number of servers and the number of racks, that is, $N^j_{\mathrm{ToR}} = N_{\mathrm{ToR}} = 1600\ \forall\, j \in N_{\mathrm{DC}}$. As many as 64 ToR switches are connected to a HOS aggregation switch using 40 Gbps links. We obtain that each data center is equipped with $N^j_{\mathrm{Aggr}} = N_{\mathrm{Aggr}} = 25\ \forall\, j \in N_{\mathrm{DC}}$ HOS aggregation nodes. Each HOS aggregation node is connected to the HOS core node inside the data center using one fiber. The HOS core switch inside a data center is equipped with 25 fiber ports for interconnecting all the HOS aggregation switches. In addition it employs 7 fiber ports for the interconnection toward the Pan-European network. Thus, it has in total 32 fiber ports. The number of fiber ports for the interconnection between a data center and the Pan-European network has been chosen according to [3], where it is reported that currently 76% of the traffic generated inside a data center is directed to a server within the same data center (internal traffic). We assume that each core node in the Pan-European network also provides edge functionality. As described before, each data center is connected to a network node of the Pan-European network using seven fibers. To ensure that the network nodes have enough capacity to support the connection toward a data center without becoming a bottleneck, we assume that each link in the network is composed of four fibers. We also assume that each HOS core node is connected to the corresponding HOS edge node using a number of fibers that is equal to the node degree. As a result, the number of fibers attached to the ith HOS edge node ($N_F^{\mathrm{Edge},i}$) is equal to the node degree. The number of fibers connected to the ith HOS core node ($N_F^{\mathrm{Core},i}$) is equal to five times the node degree (four times the node degree for the interconnection toward other HOS core nodes and one time the node degree for the interconnection toward the HOS edge node), plus seven fibers in the case that the HOS core node is directly connected to a data center. Each fiber carries 64 wavelength channels ($N_W = 64$), each of which is operated at 40 Gbps. As for the edge caching, we assume that the network nodes that are not directly connected to a data center are equipped with the same number of cache servers ($N^i_{\mathrm{CS}} = N_{\mathrm{CS}}\ \forall\, i \in N_{\mathrm{Node}}$). The network nodes that connect

data centers with the HOS core network do not comprise any cache servers.

The cache size of a HOS edge node is defined as the sum of the storage capacities, expressed in bytes, of all the video servers hosted in the node. Furthermore, we define the video content hit rate as the probability that a video request arriving at a HOS edge node determines a cache hit and thus is served using the local cache servers. The cache hit rate depends mainly on the cache size. Several studies report the cache hit rate as a function of the cache size in real networks based on the YouTube video distribution infrastructure [22,23]. The results of these studies show that high video hit rates can be achieved even with small cache sizes and that the cache hit rate exhibits a logarithmic growth as a function of the cache size. As a consequence, increasing the cache size over a certain value has a limited impact on the video content hit rate. In our simulation model, we assume that the video content popularity belongs to a Zipf distribution with a library of 2 million objects and a skew parameter equal to 0.6. These are typical assumptions for simulating a YouTube-like video content delivery service [14,15] which lead to cache hit rates consistent with those presented in [22,23]. We also assume that the size of the video contents is uniformly distributed between 100 and 500 MByte [15], with an average video size of 300 MByte, and, consequently, a library that amounts on average to 600 TByte.

The IP traffic arriving at the HOS edge nodes from legacy IP networks is modeled using a Poisson distribution. We assumed that 57% of this traffic consists of requests for video content [2]. A request for video content can be either served locally by the video cache servers in the HOS edge node if the required content is available in the cache, or can be forwarded to the original server located in one of the data centers. In our simulations we also take into account the possibility that some of the traffic that arrives at a HOS edge node is destined to another network node, that is, not to the data center. We refer to this traffic as edge-to-edge traffic. Even if it is not directly related to our analysis, the edge-to-edge traffic is important because it has an impact on the data losses and the delays as well as on the energy consumption. For the traffic generated by the servers, we implemented a more complex traffic model. According to [24], the interarrival rate distribution of the packets generated inside a data center can be modeled using a lognormal distribution. We then model the servers as finite-state machines with two states, namely the lognormal state and the video-transfer state. In the lognormal state, the servers generate IP packets with a lognormal-distributed interarrival time. The IP packets generated by the servers in the lognormal state can be addressed either to a server in the same data center, to a server in a different data center, or to a specific legacy IP network connected to a HOS edge node. When a server receives a request for video content from an edge node, it switches to the video-transfer state. In the video-transfer state the server transmits IP packets at a constant bit-rate to the requesting HOS edge node. When all the video content has been transmitted, the server automatically switches back to the state with the longnormal interarrival rate distribution.

C. Performance Metrics

The performance of the proposed integrated intra-data-center and core network architecture based on HOS is assessed in terms of energy consumption, average delay, and average data loss. The energy consumption is measured in terms of joules per bit (J/b) and is computed as the ratio between the total network power consumption in watts and the total network throughput in bits per second.

The delay is defined as the time difference between when an IP packet is generated (i.e., either by a server in a data center or a cache server in a HOS edge node or a user of a network connected to edge HOS nodes) and the time the IP packet is received (i.e., either by the destination server or the destination HOS edge node). The global average network delay is defined as the mean value of the delays over all IP packets measured during a simulation run. The IP packets that traverse the HOS network can be carried over different transport mechanisms. We refer then to the packet delay as the delay experienced by IP packets that are transmitted as optical packets through the HOS network. Similarly, the short burst delay, long burst delay, and circuit delay are the delays experienced by IP packets that are transmitted through the HOS network over a short burst, a long burst, or a circuit, respectively.

While computing the data loss rates, we assume that all the electronic switches introduce negligible losses. As a consequence, the losses in the core and intra-data-center networks may happen only in the HOS core switches. We define the packet loss rate as the ratio between the number of optical packets that are lost along a path through the HOS network and the total number of generated packets. Similarly, the short burst and the long burst loss rates are defined as the ratio between the number of lost and the number of generated short and long bursts, respectively. Circuits are established using a two-way reservation mechanism, and consequently the data transmitted over circuits do not experience any losses. However, in heavily loaded networks a circuit establishment request could be refused (i.e., blocked) by a core node. As a consequence, we define the circuit establishment failure probability as the ratio between the number of blocked and the number of generated circuits.

We evaluate the above-mentioned performance for different values of the network load. We define the load as the ratio between the total amount of traffic offered to the network by external sources (servers and legacy IP networks) and the maximum amount of traffic that can be handled by the network, that is, the network capacity.

V. Numerical Results

This section presents a performance analysis of the proposed integrated intra-data-center and core HOS network architecture with edge caching. First we comment on the benefits that a carrier cloud operator can achieve by employing the integrated HOS network instead of either a

nonintegrated HOS network or a conventional IP network. Then we present and discuss the impact of using distributed cache servers on an integrated HOS network.

A. Integrated HOS Network

To better understand the results presented in this section, in the following, we first explain the difference between the integrated and nonintegrated HOS architectures. In the nonintegrated HOS architecture, to interconnect data centers and core networks, we employ (i) HOS edge nodes (one per data center) at the core network side and (ii) HOS data-center-to-core interfaces at the data center side. These components are shown in Fig. 3. The HOS data-center-to-core interfaces perform traffic classification, conditioning, and assembling according to the data center policies, while the HOS edge nodes perform the same functions according to the policies used in the core network. The internal architecture of the HOS data-center-to-core interfaces is the same as the one of the HOS aggregation switches used inside the data center. In the integrated HOS network, there is no need for using HOS edge nodes and HOS data-center-to-core interfaces to connect data centers to the core network because both data center and core network policies are considered when processing the data center traffic in the HOS aggregation nodes. Here, the HOS edge nodes are only needed at the customer (cloud consumer) end.

In Fig. 4 we compare the energy consumption per bit as a function of the network load for the integrated HOS network, the nonintegrated HOS network, and a conventional IP network. The conventional IP network has a core and an intra-data-center network based on electronic switching. For comparative purposes, we also consider an IP core network able to put the LCs into sleep mode dynamically during idle times [18,19]. In our simulations, we assumed that all the network nodes which are not directly connected to a data center are equipped with $N_{CS} = 10$ cache servers, resulting in a total cache size of 10 TByte, which corresponds to 1/60 of the library.

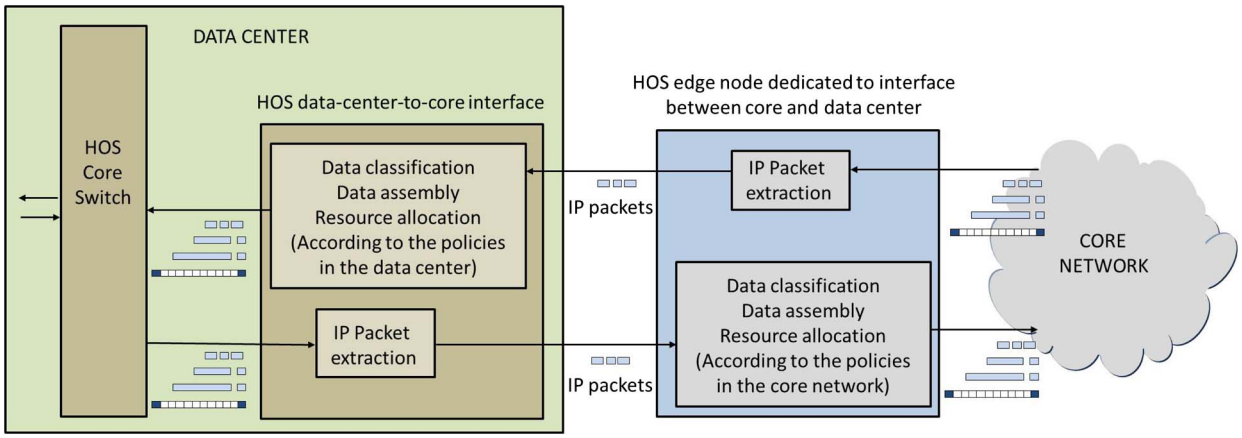

Fig. 3. Interconnection between data center and core network in the nonintegrated HOS architecture.

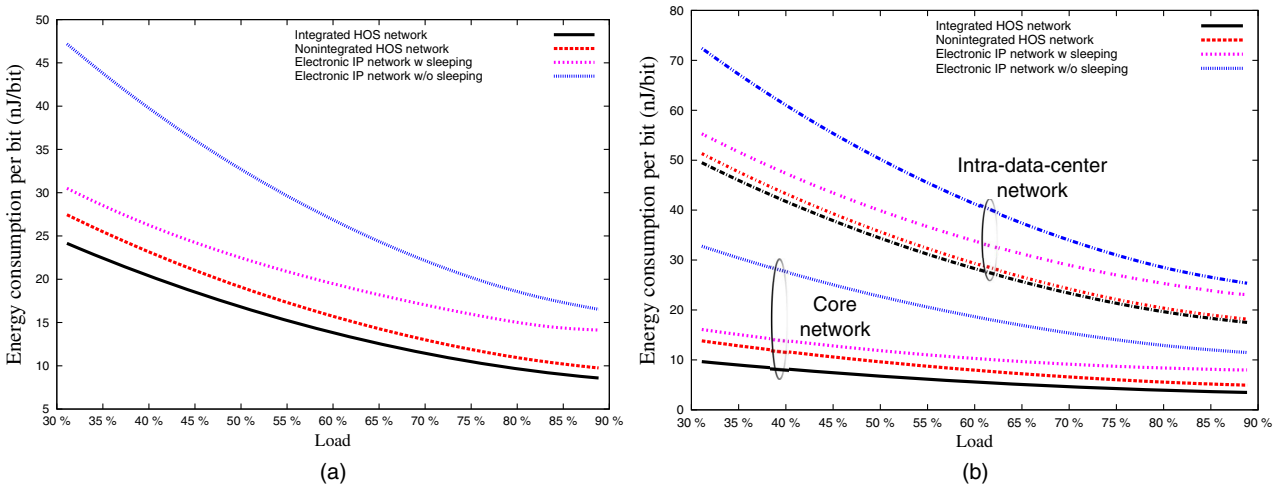

Fig. 4. Energy consumption per bit as a function of the input load. The energy consumption per bit is the ratio between the network power consumption and the network throughput. (a) Overall for core and intra-data-center networks and (b) core and intra-data-center networks shown separately.

In Fig. 4(a) we show the overall energy consumption per bit. The figure shows that by employing a sleep-based technique it is possible to achieve large energy savings with respect to current electronic IP networks, which leave all LCs in the active mode during the idle times. It is also evident that, even if a sleep-based technique is employed in an IP network, a HOS network is still able to achieve significantly lower energy consumption values. This is because the HOS networks are based on an energy-efficient optical switching technology that benefits from transmitting circuits and long bursts using an optical switch with low power consumption and relatively slow switching time while using a small number of fast optical switches for the transmission of packets and short bursts. The benefit of using HOS becomes more evident for high loads, where sleeping is not able to provide a significant improvement. Figure 4(a) also shows the improvement in energy efficiency offered by the integrated HOS network with respect to the nonintegrated HOS network. This increment in energy efficiency may seem small, but it is worth noting that at very high amounts of network traffic, such as those forecasted in [2,3] and assumed in this paper, even a reduction of a few nanojoules per bit can result in significant overall energy savings. For instance, at a network load of 35% the integrated HOS network consumes 4 nJ/b less than the nonintegrated HOS network. This translates into a total of almost 2 MW saved.

In Fig. 4(b) we show separately the energy consumption per bit in the core network and the intra-data-center networks. The energy consumption per bit of the core network is given as the ratio of the core network power consumption and the core network throughput. In the nonintegrated HOS network the power consumption of the core includes the HOS edge nodes dedicated to the interconnection toward the data centers. Similarly, the energy consumption per bit of the intra-data-center networks is given as the ratio of the total power consumption of the intra-data-center networks and the total throughput of the intra-data-center networks. In the nonintegrated HOS architecture the power consumption of the intra-data-center networks include the HOS aggregation switches dedicated to the interconnection toward the core. From the figure we have two important observations. First, core networks are more energy efficient than intra-data-center networks. In fact, according to our calculations the power consumption of the intra-data-center networks is always much higher than the power consumption of the core network for similar amounts of carried traffic. The difference is mainly coming from the ToR switches introducing an extra level of aggregation in the intra-data-center networks, which is not present in the core network [see Eqs. (2) and (8)]. The ToR switches consume a very large amount of power and dominate the power consumption of the intra-data-center networks because of the very large number of ToR switches in current high-capacity data centers based on the three-tier fat-tree network topology. This fact is more evident for the integrated HOS network where we observe that at a network load of 35% the energy consumption per bit of the core network is 5 times lower compared to the intra-data-center network. Second, the integrated approach has a higher beneficial impact on the energy consumption per bit of the core network than of the intra-data-center networks. In fact, when comparing the energy consumption of the integrated and the nonintegrated HOS networks, we observe that at a network load of 35%, the integrated approach reduces the energy consumption per bit by 30.5% of the core network and by 3.5% in the case of the intra-data-center network. This is because the additional HOS edge nodes, used in the nonintegrated HOS network to connect toward the data centers, have a strong impact on the total power consumption of the core network. This impact is higher than the impact of the additional HOS aggregation switches used inside the intra-data-center networks.

In Fig. 5 we compare the values of the average network delays as a function of the network load. Figure 5(a) shows the average delays in the integrated HOS network, while Fig. 5(b) presents the average delays in the nonintegrated HOS network. The figures demonstrate clearly that the integrated approach leads to a better delay performance and reduces the global average delays of IP packets by always more than 1 ms. In particular, the integrated approach

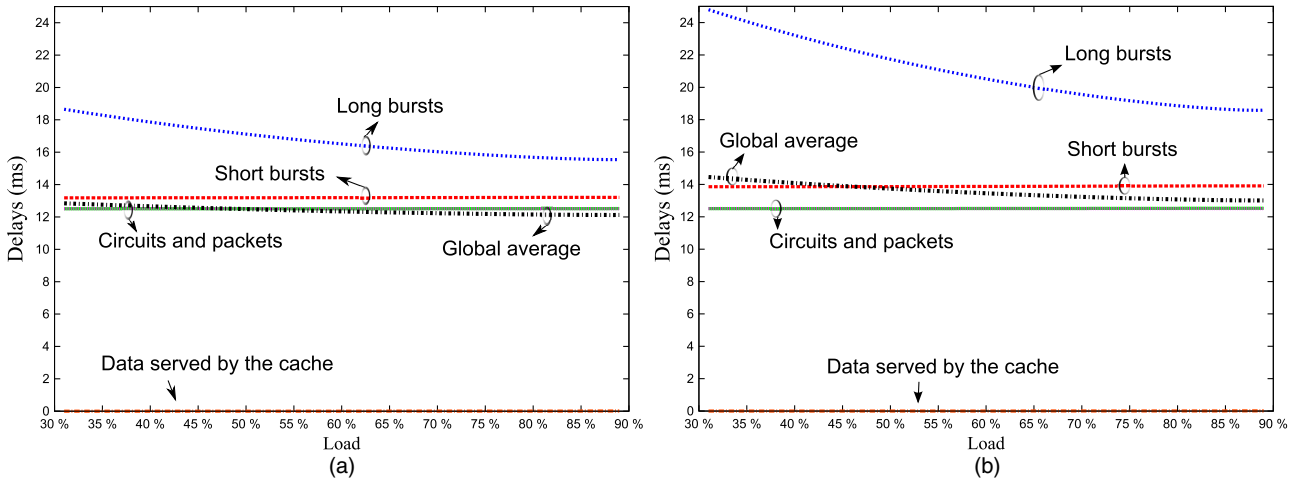

Fig. 5. Average network delays as a function of the input load for the integrated and the nonintegrated HOS networks. (a) Integrated HOS network and (b) nonintegrated HOS network.

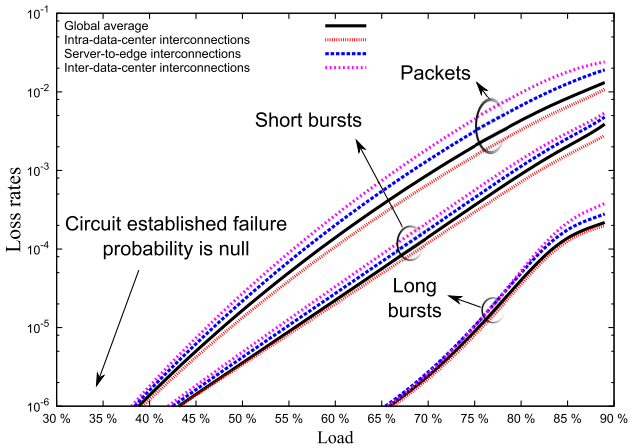

Fig. 6. Average data loss rate as a function of the input load.

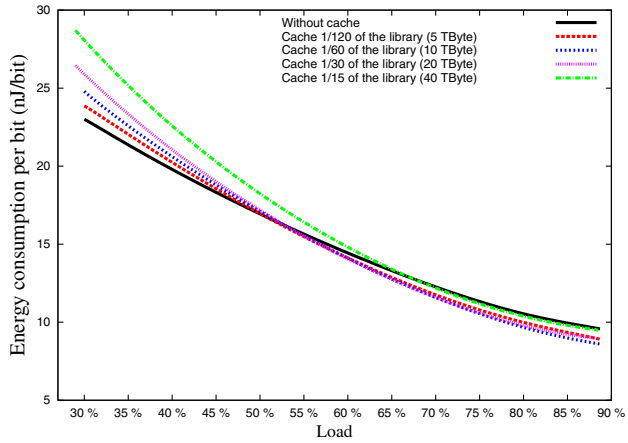

Fig. 7. Energy consumption per bit against the input load for different cache sizes.

significantly reduces the delays of IP packets transmitted over short and long bursts. This is due to the fact that bursts employ a mixed timer-length assembly algorithm [9] that may take from several hundreds of microseconds up to a few milliseconds. In the nonintegrated HOS network, the bursts must be disassembled and assembled again in the electronic interfaces between a data center and the core network, leading to a strong increase in the overall network delay.

In this paper we assume that the electronic components introduce negligible losses. As a consequence, the data loss rates in the integrated HOS network and in the nonintegrated HOS network are the same. In Fig. 6 we show the average data loss rates as a function of the network load. The optical packets are scheduled with the lowest priority, and thus they experience the highest losses. Optical bursts are scheduled *a priori* due to the offset time so that they receive a sort of prioritized handling in comparison to packets. In particular, long bursts are characterized by long offset times and show loss rates almost three orders of magnitude lower than packets and almost two orders of magnitude lower than short bursts. Finally, circuits are scheduled with the highest priority and achieve a lossless operation and negligible establishment failure probabilities in our simulations. To understand where in the network we observe the highest losses, we plot in Fig. 6 the average loss rates in the intra-data-center, inter-data-center, and server-to-edge interconnections. We observe that the average loss rates in inter-data-center interconnections are always the highest. This is because in the inter-data-center interconnections the data needs to cross on average the highest number of HOS core switches (in both the HOS core network and intra-data-center network). The lowest average loss rates are instead achieved by the intra-data-center interconnections where data always have to cross a single HOS core switch inside the data center.

B. Impact of Edge Caching

In Fig. 7 we show the energy consumption per bit of the integrated HOS network against the network load and for different values of the cache size. To vary the cache size we change the number of cache servers per HOS edge node, that is, the value of N_{CS}. We always assume that the network nodes connected to a data center are not equipped with local cache servers. To understand the results shown in Fig. 7 it should be noted that we do not consider dynamic switching-off of the electronic LCs, and consequently the energy consumption of the electronic components is independent on the network load. Only the power consumption of the optical switching fabric of the HOS core nodes, that is, P_{OSF}, changes with the network load. Furthermore, it should be noted that the energy consumption per bit is defined as the ratio between the network power consumption given in watts and the network throughput given in bits per second. Figure 7 shows that at low and moderate loads, the higher the cache size, the higher the energy consumption per bit. In fact, in our simulations, the increase in the storage energy consumption introduced by the distributed cache servers (P_{Cache}) is always higher than the reduction of the transport energy that is obtained by switching off the unused optical switch ports of the HOS core nodes. When increasing the load, we observe that the larger the cache size, the faster the decrement of the energy consumption per bit. This is due to the fact that increasing the number of distributed cache servers reduces the average data loss rates in the network. The larger the cache size, the higher the network throughput, especially at high loads. However, the network throughput does not increase linearly with the cache size. In fact, as shown in [22,23], the network throughput increases in a log-like way with the increase in the cache size. This means that the network throughput becomes saturated when increasing the cache size over a certain value. On the other hand, increasing the cache size leads to an almost linear increase of the storage power consumption. As a consequence, at high loads we observe that there is a trade-off between cache size and energy consumption per bit. In our simulations, when the load is higher than 50%, the best results in terms of energy consumption are achieved using 1/60 of the size of the library, that is, setting $N_{\mathrm{CS}} = 10$.

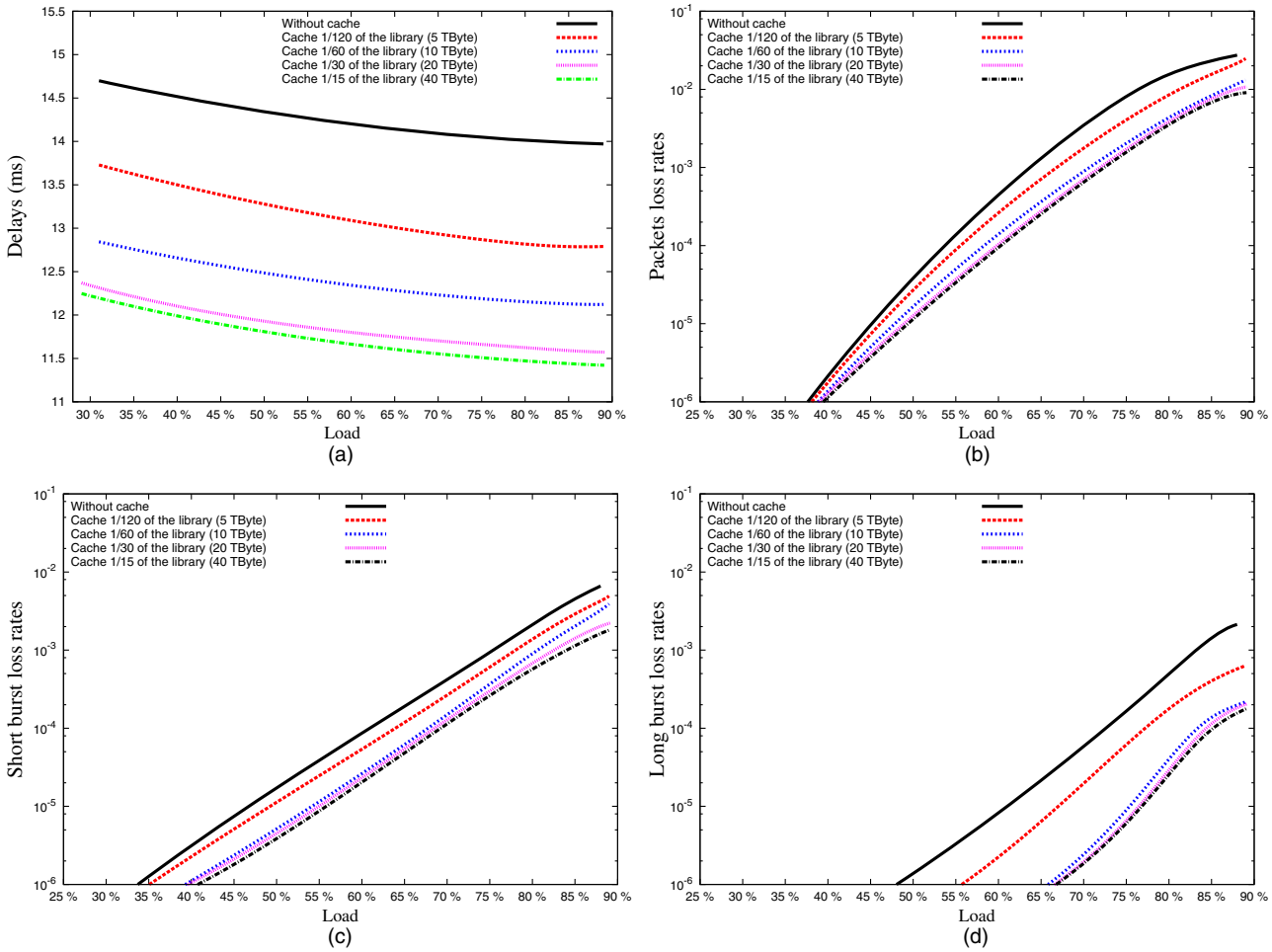

Fig. 8. Average delays and average data loss rates as a function of the input load for different values of the cache size. (a) Average network delays, (b) packet loss rates, (c) short burst loss rates, and (d) long bursts loss rates.

In Fig. 8(a) we present the global average network delay as a function of the network load for different values of the cache size. The figure highlights that the larger the cache size, the lower the global average network delay. In particular, increasing the cache size from 0 to 1/30 of the size of the library (i.e., from 0 to 20 TByte) leads to a reduction of the global average delay in the network by about 2 ms. A further increase of the cache size from 1/30 to 1/15 of the size of the library (i.e., from 20 to 40 TByte) has a very limited impact on the global average network delays.

Finally, in Figs. 8(b)–8(d) we show the average loss rates of packets, short bursts, and long bursts as a function of the network load for different values of the cache size. The circuit establishment failure probability is always null in the considered configurations. The figures show that the larger the cache size, the lower the average loss rates. This is due to the fact that increasing the cache size keeps the traffic more local, which corresponds to a higher amount of requests from the end users served by the cache servers. This leads to a reduction of the traffic in the core and in the intra-data-center networks and consequently to lower loss ratios. Figure 8 also shows that by increasing the cache size from 0 to 1/60 of the size of the library (i.e., from 0 to 10 TByte) we achieve a high reduction in the loss rates, while increasing the cache size over 1/60 of the size of the library (i.e., over 10 TByte) has a very limited impact on the loss rates.

VI. Conclusions

In this paper we have proposed a unified network architecture that provides both intra-data-center and inter-data-center connectivity together with interconnection toward legacy IP networks. This architecture is tailored for the future carrier cloud operators running both the data centers and the core network. The architecture is referred to as an integrated core and intra-data-center network and is based on the HOS technology. The main advantage of the integration of core and intra-data-center networks in a single infrastructure comes from avoiding electronic interfaces between the data centers and the core network. We evaluated the energy consumption along with the delay and loss performance of the integrated HOS network and made extensive comparisons with respect to a nonintegrated HOS solution and a conventional IP

network based on electronic switching. We conclude that the integrated HOS network achieves by far the highest energy efficiency. Furthermore, we demonstrated that the integrated HOS network reduces considerably the average network delays with respect to a nonintegrated HOS solution. As a consequence, we conclude that the integrated HOS network is well suited for application in carrier clouds.

Furthermore, we studied the impact of distributed video cache servers on the energy consumption as well as the delay and loss performance of the integrated HOS network. The existing literature on this topic only takes into account conventional core and intra-data-center networks based on IP electronic switching, which are characterized by low energy efficiency. The aim of this study is to identify whether a carrier cloud operator that relies on the integrated HOS network concept could increase energy efficiency by employing edge caching. Therefore, we extended HOS edge node architecture to include cache servers and content trackers. The content trackers interact with the HOS control plane for updating the servers and processing incoming video requests. We also developed a novel analytical model for evaluating the energy consumed by the cache. According to the results we conclude that to achieve both low delay and data loss as well as high energy efficiency in an integrated HOS network, a careful dimensioning of the cache size is needed. In particular, at low and moderate loads we observed the highest energy efficiency is achieved in the case without any edge caching. Furthermore, our analysis also leads to the following general conclusion: when deciding to upgrade the traditional electronic switching-based network to a more-energy efficient one, operators have to reconsider their edge caching strategy in order to achieve the best network performance.

References

[1] "SMART2020: Enabling the Low Carbon Economy in the Information Age," The Climate Group, Global eSustainability Initiative, Tech. Rep., 2008 [Online]. Available: www.smart2020.org.

[2] "Cisco Visual Networking Index: Forecast and Methodology, 2012–2017," Cisco White Paper, May 2013.

[3] "Cisco Global Cloud Index: Forecast and Methodology, 2011–2016," Cisco White Paper, May 2012.

[4] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy efficiency in telecom optical networks," *IEEE Commun. Surv. Tutorials*, vol. 12, no. 4, pp. 441–458, Fourth Quarter 2010.

[5] R. S. Tucker, "Green optical communications part II: Energy limitations in networks," *IEEE J. Sel. Top. Quantum Electron.*, vol. 17, no. 2, pp. 245–260, Mar./Apr. 2011.

[6] "Where does power go?" GreenDataProject, 2008 [Online]. Available: http://www.greendataproject.org.

[7] C. Kachris and I. Tomkos, "A survey on optical interconnects for data centers," *IEEE Commun. Surv. Tutorials*, vol. 14, no. 4, pp. 1021–1036, Fourth Quarter 2012.

[8] R. Veisllari, S. Bjornstad, and D. Hjelme, "Experimental demonstration of high throughput, ultra-low delay variation packet/circuit fusion network," *Electron. Lett.*, vol. 49, no. 2, pp. 141–143, Jan. 2013.

[9] M. Fiorani, M. Casoni, and S. Aleksic, "Hybrid optical switching for energy-efficiency and QoS differentiation in core networks," *J. Opt. Commun. Netw.*, vol. 5, no. 5, pp. 484–497, May 2013.

[10] O. Liboiron-Ladouceur, I. Cerutti, P. Raponi, N. Andriolli, and P. Castoldi, "Energy-efficient design of a scalable optical multiplane interconnection architecture," *IEEE J. Sel. Top. Quantum Electron.*, vol. 17, no. 2, pp. 377–383, Mar./Apr. 2011.

[11] M. Fiorani, S. Aleksic, and M. Casoni, "Hybrid optical switching for data center networks," *J. Electr. Comput. Eng.*, vol. 2014, 139213, 2014.

[12] J. Baliga, R. Ayre, K. Hinton, and R. S. Tucker, "Architectures for energy-efficient IPTV networks," in *Optical Fiber Communication Conf. (OFC)*, 2009, paper OThQ5.

[13] C. Jayasundara, A. Nirmalathas, E. Wong, and C. Chan, "Energy efficient content distribution for VoD services," in *Optical Fiber Communication Conf. (OFC)*, 2011, paper OWR3.

[14] C. Chan, E. Wong, A. Nirmalathas, A. Gygax, and C. Leckie, "Energy efficiency of on-demand video caching systems and user behavior," *Opt. Express*, vol. 19, no. 26, pp. B260–B269, Dec. 2011.

[15] N. Osman, T. El-Gorashi, and J. Elmirghani, "The impact of content popularity distribution on energy efficient caching," in *Proc. Int. Conf. on Transparent Optical Networks (ICTON)*, 2013, pp. 1–6.

[16] D. Cai and S. Natarajan, "The evolution of the carrier cloud networking," in *Proc. IEEE Symp. on Service-Oriented System Engineering (SOSE)*, 2012, pp. 286–291.

[17] A. Autenrieth, J. Elbers, P. Kaczmarek, and P. Kostecki, "Cloud orchestration with SDN/OpenFlow in carrier transport networks," in *Proc. Int. Conf. on Transparent Optical Networks (ICTON)*, 2013, pp. 1–4.

[18] F. Idzikowski, S. Orlowski, C. Raack, H. Woesner, and A. Wolisz, "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," *Proc. Conf. on Optical Network Design and Modeling (ONDM)*, 2010, pp. 1–6.

[19] S. Nedevschi, L. Popa, G. Iannaccone, S. Ratnasamy, and D. Wetherall, "Reducing network energy consumption via sleeping and rate-adaptation," in *Proc. USENIX Symp. on Networked Systems Design and Implementation*, 2008, pp. 323–336.

[20] A. Betker, C. Gerlach, R. Hulsermann, M. Jager, M. Barry, S. Bodamer, J. Spath, C. Gauger, and M. Kohn, "Reference transport network scenarios," MultiTeraNet Report, July 2003.

[21] "Connectivity solutions for the evolving data center," Emulex White Paper, May 2011.

[22] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Characteristics of YouTube network traffic at a campus network: Measurements, models, and implications," *Comput. Netw.*, vol. 53, no. 4, pp. 501–514, Mar. 2009.

[23] L. Braun, A. Klein, G. Carle, H. Reiser, and J. Eisl, "Analyzing caching benefits for YouTube traffic in edge networks: A measurement-based evaluation," in *Proc. IEEE Network Operations and Management Symp. (NOMS)*, 2012, pp. 311–318.

[24] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proc. Internet Measurement Conf. (IMC)*, 2010, pp. 267–280.