Performance Evaluation of an All-Optical OCS/OPS-Based Network for Intra-Data Center Connectivity Services

Albert Pagès, Sergio Jiménez, Jordi Perelló, and Salvatore Spadaro, *Member, IEEE*
Advanced Broadband Communications Centre (CCABA), Universitat Politècnica de Catalunya (UPC)
Jordi Girona 1-3, 08034 Barcelona, Spain
Email: {albertpages, sergio.jimenez, spadaro}@tsc.upc.edu, perello@ac.upc.edu

ABSTRACT
As the size of Data Centers (DCs) and the amount of information managed by them grows, nowadays electronic-based intra-Data Center Networks (DCNs) start to become a limiting factor, not providing enough scalability nor flexibility to support highly dynamic traffic profiles. The LIGHTNESS intra-DCN solution, based on a hybrid all-optical OCS/OPS architecture, aims to overcome such limitations, providing an architecture suitable for the management of both long- and short-lived traffic flows. In this paper, we evaluate the performance of such a hybrid architecture in terms of blocking probability for various traffic conditions, highlighting its benefits against pure OCS or OPS networks.
Keywords: intra-DCNs, OPS, OCS.

1. INTRODUCTION

In the last few years, emerging new applications such as video streaming, cloud computing and social networks have driven an exponential increase on the amount of data that DC have to process. Additionally, such applications require high interaction between servers inside a DC. Some forecasts predict that global IP traffic in DCs will increase by almost a factor of 2 in the following three years, reaching a yearly traffic of 7.7 Zettabytes by 2017 [1]. Moreover, the same forecasts predict that in such period of time, around 76% of the global traffic handled by DCs will be intra-DC. Such substantial growth will push the limits in scalability, performance and power consumption of currently deployed intra-DCNs.

Nowadays DCs are built around commodity servers organized in racks, where a Top-of-the-Rack (ToR) electronic switch enables the communication between servers of the same rack and those from other racks. These ToRs are usually connected between them thanks a multi-tier electronic-based network fabric like the one depicted in Fig. 1 [2]. Unfortunately, such intra-DCN architecture suffers from serious issues such as high latencies, bandwidth limitations and large power consumptions; it cannot thus cope with the expected intra-DC traffic growths. For these reasons, it is mandatory to introduce newer network architectures capable to achieve the desired flexibility and scalability so as to successfully manage such traffic growths.

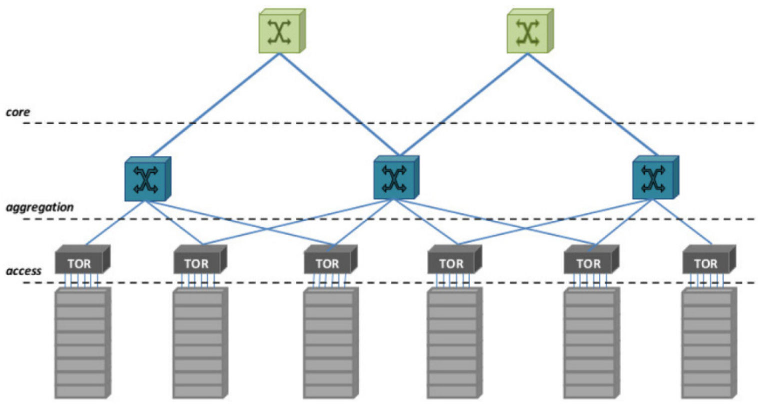

Figure 1. Current Data Center network architecture.

A very promising approach that has already spanned multiple research initiatives and is still attracting the effort of the research community is the introduction of optical technologies in the intra-DCNs [3]. By substituting currently deployed Ethernet- or Infiniband-based intra-DCN with all-optical networks, it will be possible to overcome the limitations of such architectures and shape the future DC scenarios. Thanks to Dense Wavelength Division Multiplexing (DWDM), it is possible to offer several optical channels per fiber link, each one of them possibly offering transmission rates of 100 G/s and more, resulting in a bandwidth of several terabits per second per fiber link. Additionally, thanks to transparent optical networks, it will be possible to reduce the power consumption of the intra-DCN fabric, due to the fact that electrical processing is no longer needed in all intermediate nodes in the network, effectively cutting down the overall OPEX of the DC. Moreover, the

978-1-4799-5601-2/14/$31.00

introduction of optics in the DCs will allow flattening the actual fat-tree architecture by eliminating the aggregation level, allowing also reductions on the needed CAPEX.

Following these trends, we introduce the European Framework Programme 7 (FP7) LIGHTNESS project [4], which aims to design, implement and experimentally demonstrate a high-performance intra-DCN based on the introduction of innovative optical solutions. In the following section, we detail the general architecture of the LIGTNESS intra-DCN solution. Then, the benefits of introducing hybrid optical solutions are highlighted.

2. THE LIGHTNESS DCN ARCHITECTURE

As commented previously, several research initiatives are currently active to re-design nowadays multi-tier intra-DCN architectures by means of introducing optical technologies inside the DCs. In this respect, the EU FP7 LIGHNTESS project has defined a novel flat architecture for intra-DCNs based on the integration of both Optical Circuit Switching (OCS) and Optical Packet Switching (OPS) technologies, along with a Software Defined Network (SDN) control framework [5]. Figure 2 depicts the overall LIGHTNESS DCN architecture, including the hybrid OCS/OPS data plane and the SDN control plane.

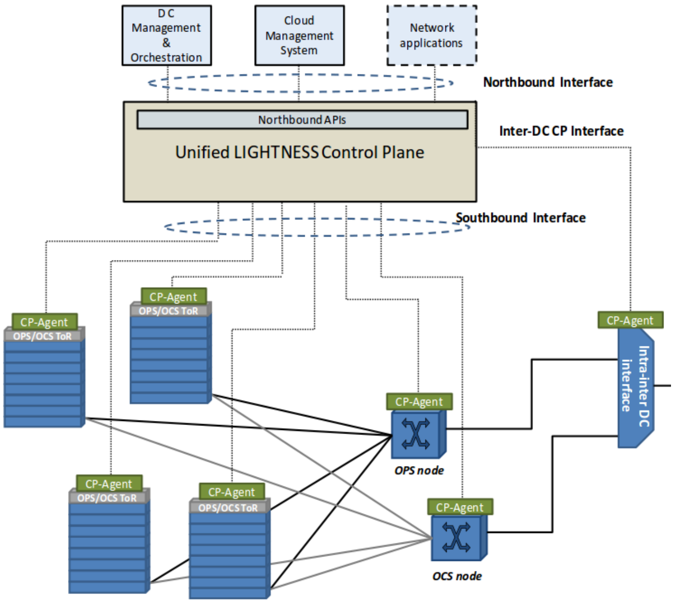

Figure 2. Overall LIGHTNESS DCN architecture.

In DC environments, applications generating long-lived traffic flows between servers coexist with applications generating short-lived traffic flows between servers. Furthermore, there is also heterogeneity in regards of the amount of bandwidth requested by traffic flows, with flows requiring high amounts of bandwidth (*e.g.* backups) and flows requiring less demanding bandwidth (*e.g.* search queries). Due to this heterogeneous traffic profile, it is really challenging to efficiently serve all traffic demands employing a single switching technology. For this reason, the LIGHTNESS data plane relies on both OCS and OPS technologies so as to properly serve connection requests between servers depending on their specific characteristics and requirements.

Specifically, OCS will be employed to serve long-lived data flows, since the set-up time required to establish an optical circuit may be not worth for serving short-lived data flows. On the other hand, OPS will be employed to serve short-lived traffic flows in order to target low latency flow requirements. As for the distinction between high and low bandwidth demands, one possibility is to map the most demanding, bandwidth speaking, traffic flows to OCS, thanks to their capacity to almost fill the entire bandwidth provided by a single wavelength. Conversely, less demanding traffic flows will be mapped to OPS, since its statistical multiplexing capabilities allows the sharing of the bandwidth of a single wavelength among lower bit-rate traffic flows. All the servers of a rack are connected to the hybrid OCS/OPS optical fabric via the ToR switch, which will be responsible to perform traffic classification and aggregation.

The unified SDN-based control plane, along with the management and orchestration modules, is designed to overcome the limitations of current DCN control and management solutions. Basically, the unified control plane implements all the necessary functionalities and mechanisms so as to allow dynamic setup, recovery and optimization of high-capacity connectivity services among servers/racks. Moreover, thanks to the communication between the control plane and the management modules, it is possible to take into account both application requirements and DCN status (including IT resources), allowing for a joint optimization of the DC resources.

In the following section, we will evaluate the presented intra-DCN architecture when faced with dynamic connectivity requests between servers, benchmarking it against optical networks that only employ one of the two considered switching technologies, either OCS or OPS. With this, we aim to highlight the superior performance and scalability of the proposed hybrid solution.

3. PERFORMANCE EVALUATION

3.1 Scenario Description

Before proceeding with the actual performance evaluation results and discussion, let us first disclose the intended evaluation scenario. Following the network architecture depicted in Fig. 2, our evaluation will focus on network architecture that comprise both OCS and OPS switching technologies. To this end, we define a DC scenario composed of multiple servers arranged in racks that are connected to both a OCS switch and a OPS switch via a ToR switch, one for each of the racks. As we are aiming to evaluate the network performance, without loss of generality, we will assume that there is enough capacity to connect the servers to the ToRs and all connection requests are between ToRs. Additionally, we will assume that all the connection requests between servers are from servers that belong to different racks. Lastly, since we are targeting to tests transparent optical services, the wavelength continuity must be respected along the end-to-end path, either in OCS or OPS optical flows.

As for the characteristics of the connections, we assume that all connections request for a specific bandwidth in terms of percentage capacity of a single wavelength as well as a specific Quality of Service (QoS). The requested QoS relates to the average packet loss ratio that the traffic flow can support. This QoS value is imposed due to the packet contention that may happen in the OPS switch if traffic flows are served employing said technology. Contention happens when two or more optical packets from different input ports simultaneously need to be sent to the same output port. In this situation, since optical signals would collide, only one of the packets can be actually sent while the rest have to be dropped.

This situation happens due to the inexistence of optical buffer to store the packets that would allow performing contention management and overcome such limitation. Additionally, this phenomenon increases exponentially with the traffic load supported by the wavelengths because then the collision probability is higher. Since there exists a correlation between the number of ports of an OPS switch, the wavelengths load and the packet loss ratio due to the contention, it is necessary to limit the actual load carried by any wavelength employed by OPS flows to the most restricting QoS value of the flows circulating by that particular wavelength. As a consequence, it may be necessary to split a single connectivity request among multiple OPS flows so as to respect the QoS of any traffic flow circulating in any of the wavelengths. Nevertheless, it is still possible to share a wavelength between multiple OPS flows, thanks to the statistical multiplexing property, as long as all QoS are respected.

On the other hand, if a demand is served employing OCS, the packet loss ratio is no longer relevant and the QoS of the flow is satisfied. This is due to the fact that an optical circuit, once established, is exclusively reserved for the traffic flow employing that particular circuit, so no contention can happen in this situation.

With this context, we will evaluate the performance in terms of blocking probability of incoming connection requests between ToRs in the DCN. Due to the fact that a connection request can be mapped to OCs or OPS depending on its characteristic, it is necessary to perform some decisions in this regard. To this end, we define the parameter *T* that will act as a threshold in order to decide when a particular demand is mapped to OCS or OPS. The decision criterion is as follows: if the bandwidth requested by a connection is lower or equal than *T*, the demand will be served employing OPS; if it is greater than *T*, then it will be served employing OCS. It can be appreciated that the actual value of *T* plays a critical role on the overall performance of the proposed DCN architecture. To this end, we will show how the performance of the network evolves with the value of *T*. Additionally, if we set the value of *T* in terms of percentage use of a wavelength, it can be seen that *T*=0% reflects the pure OCS situation, while *T*=100% reflects the pure OPS situation. The following section presents the details of the actual test scenario and the obtained results.

3.2 Results and Discussion

In order to test the proposed architecture, we have performed a series of simulations. Particularly, we assume a network topology composed of 16 racks, with their corresponding ToR connected to both an OCS switch and an OPS switch through a fiber link. Each fiber link is equipped with 64 wavelengths. We have generated 10^6 random connection requests arriving to the network, following a Poisson process with exponentially distributed Holding Times (HTs) and Inter-Arrival Times (IATs), so the network load is computed as HT/IAT. Connections are characterized by their source node R_s and their destination node R_d, which are chosen at random following a uniform distribution among the ToRs in the network such as $R_s \neq R_d$; by their requested bandwidth, which is randomly chosen between 1% and 100%, in steps of 1, in terms of wavelength capacity; and their required QoS in terms of packet loss ratio, which is chosen randomly between $\{10^{-6}, 10^{-3}, 10^{-2}\}$, for which a 16×16 OPS switch corresponds to maximum allows loads per wavelengths of {35%, 40%, 44%}. In the following figure, we show how the blocking probability of the connections evolves as a function of the load and for various values of *T*.

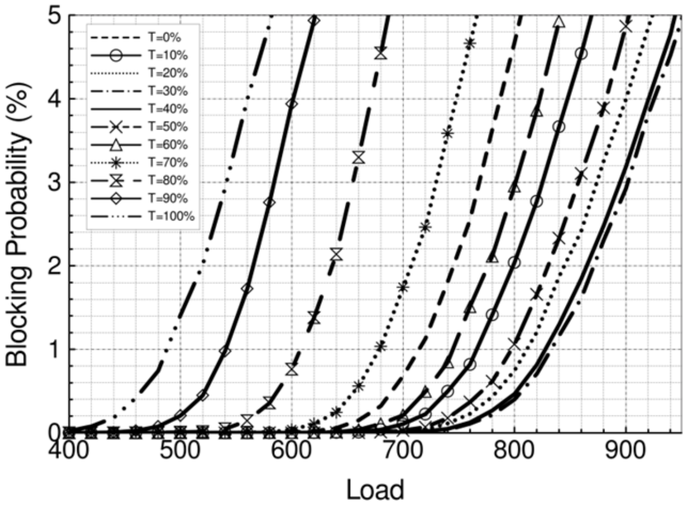

Figure 3. Blocking probability as a function of the load for multiple values of T.

By inspecting the figure it can be noticed that threshold 30% provides the best (lowest) blocking probability. It can also be noticed that an only OCS network (threshold 0%) would yield a lower blocking probability than thresholds from 100% (only OPS) to 70% but higher than thresholds from 60% to 10%. An only OPS solution (threshold 100%) would yield the highest blocking probability. Nevertheless, with a value of $T = 30\%$, the proposed hybrid architecture performs better than only considering one of the two switching technologies.

Additionally, it can be appreciated that the value of *T* that provides the lowest blocking figures is very close to the most restricting value of QoS a demand can request (35%). This happens because, if a demand that is mapped to OPS requests more bandwidth than their requested QoS, it will be necessary to split the demand into multiple OPS flows so as to guarantee the QoS for the demand, resulting in more resources occupied, hence, higher blocking probability.

4. CONCLUSIONS AND FUTURE WORK

We have presented the LIGHTNESS architecture as a potential enabling solution for future intra-Data Center communications. We have shown that the proposed hybrid OCS/OPS-based optical network performs better in terms of blocking probability than considering a network with a single switching technology, either OCS or OPS. Particularly, if a proper mapping to OCS or OPS of the data flows is performed, relative gains in terms of supported load for a BP of 1% around 75% and 18% can be observed when compared to pure OPS and pure OCS networks respectively.

As for future work, we devote our effort investigating techniques to find the optimal network topology for serving a particular demand set. Additionally, we will evaluate the proposed architecture for the case where Virtual Data Center (VDC) requests are considered, not only VM pairs. To this end, we will investigate the necessary optical resources needed to properly satisfy a given VDC request set, evaluating also the cost and power consumption of the proposed solution against current electrical-based networks.

ACKNOWLEDGEMENTS

The work in this paper has been supported by the Spanish Science Ministry through the ELASTIC project (TEC2011-27310).

REFERENCES

[1] Cisco Global Cloud Index 2012, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html.

[2] A. Predieri *et al.*: The LIGHTNESS project, Deliverable 2.2 – Architecture design, http://www.ict-lightness.eu/wp-content/uploads/2013/10/Architecture-Design.pdf, November 2013.

[3] C. Kachris and I. Tomkos: A survey on optical interconnects for data centers, *IEEE Communications Surveys & Tutorials*, vol. 14, num. 4, pp. 1021-1036, Jan. 2012.

[4] The LIGHTNESS project, http://www.ict-lightness.eu

[5] J. Perelló *et al.*: All-optical packet/circuit switching-based data center network for enhanced scalability, latency, and throughput, *IEEE Network*, vol. 27, no. 6, pp. 14-22, Dec. 2013.