White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/99512/

Version: Accepted Version

---

**Proceedings Paper:**

Hammadi, AA, El-Gorashi, TEH, Musa, MOI et al. (2016) Server-Centric PON Data Center Architecture. In: 18th International Conference on Transparent Optical Networks (ICTON 2016). ICTON 2016, 10-14 Jul 2016, Trento, Italy. IEEE. ISBN: 9781509014675. EISSN: 2161-2064.

https://doi.org/10.1109/ICTON.2016.7550695

---

© 2016, IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS University of Sheffield UNIVERSITY of York

# Server-Centric PON Data Center Architecture

**Ali Hammadi, Taisir E.H. El-Gorashi, Mohamed O.I. Musa, and Jaafar M.H. Elmirghani**
*School of Electronic & Electrical Engineering, University of Leeds, LS2 9JT, United Kingdom*

**Abstract**

Over the last decade, the evolution of data center architecture designs has been mainly driven by the ever increasing bandwidth demands, high power consumption and cost. With all these in mind, a significant potential to improve bandwidth capacity and reduce power consumption and cost can be achieved by introducing PONs in the design of the networking fabric infrastructure in data centers. This work presents a novel server-centric PON design for future cloud data center architecture. We avoided the use of power hungry devices such as switches and tuneable lasers and encouraged the use of low power passive optical backplanes and PONs to facilitate intra and inter rack communication. We also tackle the problem of resource provisioning optimization and present our MILP model results for energy efficient routing and resource provisioning within the PON cell. We optimized the selection of hosting servers, routing paths and relay servers to achieve efficient resource utilization reaching 95% and optimum saving in energy consumption reaching 59%.

**Keywords**: passive optical network (PON), data centre, energy efficiency, resource provisioning, optimization.

## 1. Introduction

Despite the radical evolution of data center infrastructure designs in the last decade, current designs still suffer from many limitations [1]. Infrastructure designs that can provide scalability, energy efficiency, and low cost to sustain the ever growing large-scale cloud services are a hot area for research. Advances in optical technologies and its proven performance in access networks have attracted many researchers to introduce PONs in the infrastructure design of future data centers. Here we propose a new architecture based on PONs, with the rationale of improving the data centre energy efficiency. The proposed architecture capitalises on the PON's capability to supply on demand elastic bandwidth to interconnect the data centre elements that provide processing power, memory, and storage. The use of PONs can also address many issues that have existed in conventional data center architectures for a while now, such as switch oversubscription and load balancing.

In our previous work we investigated energy efficiency for core networks with data centers and clouds [2-9]. In [10], we introduced five novel designs for PON deployment in future cloud data centers to handle intra- and inter-rack communications. In [11], we proposed an AWGR-based PON architecture and have shown that energy savings of 45% and 80% can be achieved compared to the Fat-Tree [12] and BCube [13] architectures, respectively. The main drawback of this design is its high deployment cost as all servers are equipped with tuneable transceivers. In this paper, we present another novel scalable, high capacity, energy efficient PON data center design that eliminates the use of costly tuneable lasers. We also study its routing and resource provisioning. We present our results obtained through a developed Mixed Integer Linear Programming (MILP) model that minimizes power consumption by optimizing the location of selected servers to route and provision physical resources of CPU and memory to clients.

The remainder of this paper is structured as follows: In Section 2, we describe our proposed PON server centric data center architecture. In Section 3, we present our results for resource provisioning through our optimization model. Finally we conclude the paper in Section 4.

## 2. Architecture of the Server-Centric PON Data Center

In this section, we describe the architecture of our proposed server centric data center; relying mostly on PONs to facilitate inter- and intra-rack communication. The design, depicted in Figure 1, eliminates the need for costly tuneable lasers and facilitates high speed interconnection among racks within the PON cells. In a PON cell, each rack is divided into 4 groups, each of which can host 8 servers connected by a TDM star coupler. Three of the 4 groups are connected to the other three racks, and one group connects the rack to the OLT switch. Inter-rack communication between servers that do not belong to groups with a direct connection is established by using one of the servers of the group with a direct connection with the rack of the destination server. Relay server selection can be based on servers' utilization or traffic load within the group. Inter-cell communication can be established through servers in groups with direct connection with the OLT switch via AWGR. The OLT switch capacity can accommodate 16 Access Modules (AM), each of which can have 16 ports with transmission rate of 10Gb/s each (e.g., XGPON). Assuming a single port to connect a PON cell with a total of 128 servers, one OLT chassis can accommodate a total of 32,768 servers. The architecture can be scaled up by adding more chassis.

For Intra-rack communication, we have proposed two solutions; one deploys a Fibre Brag Grating (FBG) after the star coupler connecting the servers in each rack to reflect a dedicated wavelength assigned for intra-rack communication. This option requires either equipping servers with a multi-wavelength (MW) transceiver or introducing OFDM technology to allow a single transceiver to generate multiple carriers, one for intra-rack communication and another for connections to the OLT or other racks. The deployment cost of the FBG with MW or OFDM transceivers is expensive and introduces high complexity in the MAC protocol for coordinating the arbitration of channel access. Therefore; we encourage the use of the second option; the terabit capacity passive polymer optical backplane proposed by Cambridge [14]. This technology is shown in Figure 2, it employs a passive backplane with multimode polymer waveguides and can provide non-blocking full mesh connectivity with 10 Gb/s rates per waveguide, exhibiting a total capacity of 1 Tb/s.

## 3. Results

In this section we evaluate the power consumption and servers' utilization (where servers have forwarding and processing VMs) in the proposed PON data center design depicted in Figure 1. We model a PON cell with 18 servers distributed in three racks each hosting 6 servers. Each rack is divided into 3 groups where each group consists of 2 servers. Table I summarizes the input parameters of the model. In this work, we assumed that 5% of a server's CPU resources is needed to process a relayed request for header inspection and forwarding.

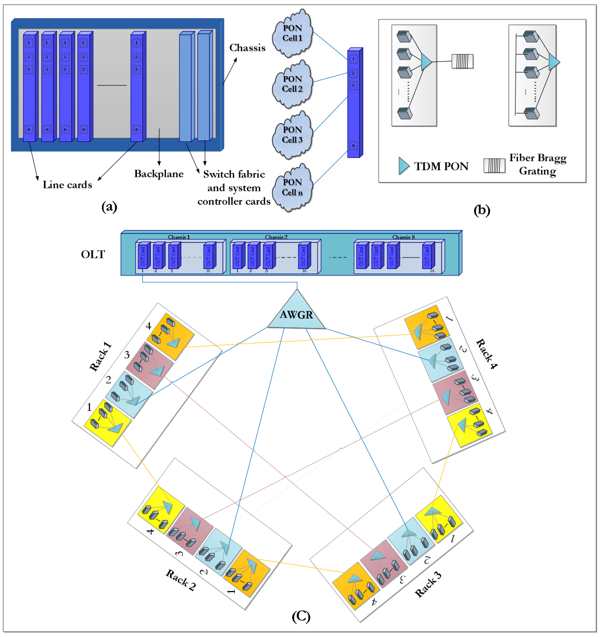

Figure. 1. A possible PON deployment in a data center (a) An OLT chassis with 16 AM, (b) Technologies for intra-rack communication; FBG and optical backplane, (c) Proposed PON Cell design with no tuneable lasers

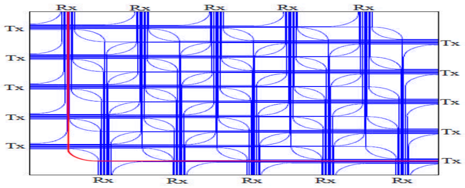

Figure.2 Schematic of 10x10 passive polymer optical backplane [14]

### TABLE I: INPUT DATA FOR THE MODEL

| | |
|---|---|
| Link capacity | 10 Gb/s |
| Power consumption for idle servers | 201 W [15] |
| Maximum power consumption for servers | 301 W [15] |
| Server's utilization for processing a relayed request (CPU cycles) | 5% |
| Clients processing requirements in CPU GHz | 500-2000 |
| Clients memory requirements in MB | 500-2000 |
| Number of clients (VMs) | 20, 30, 40, 50 |
| Server's processing capacity in GHz | 2.5 |
| Server's memory (RAM) in GB | 8 |

In the proposed design servers are used to store and process data as well as participate in traffic forwarding. For resource provisioning in the proposed design, selection of host servers is of premium importance as the scheduler attempts not only to slice servers' resources to be shared by multiple VMs, but also to reduce number of servers relaying requests optimally. For comparison purposes and to understand the behaviour of the MILP model in minimizing the power consumption, we have modelled the described network with two objectives; one for energy minimization (EA) and one for only provisioning all VMs without targeting energy saving (NEA).

The obtained results from the MILP as depicted in Figures 3 and 4 show the relation between the numbers of servers required to provision demanding clients and the location of selected servers to process and relay traffic within the PON cell. In EA MILP model, the selection of host servers amounts to server consolidation and therefore results in the minimization of the number of intermediate relays in the path aiming at minimizing the total power consumption.

Location and utilization of assigned relay servers are depicted in Figures 3 and 4 for EA and NEA models, respectively. The results show that EA model avoids the use of relay servers as far as possible and efficintly utilizes servers' resources. While the NEA model randomly selects host servers and routes. Random assignment results in lower utilization of servers resources (CPU and memory) and results in the selection of long paths with multiple intermediates servers. This clearly explains the increase in power consumption and high number of servers invloved in the routing and provisioning of resources for the NEA model compared to the EA model. Figure 6 shows the average utilization of servers for the different sets of examined VMs. On average, servers' utilization of the EA and NEA model are 95% and 60%, respectively. Power consumption of the EA and NEA models are shown in Figure 5 for the different sets of examined VMs. The results in Figure 5 show an average power saving of 35% for the examined number of VMs.

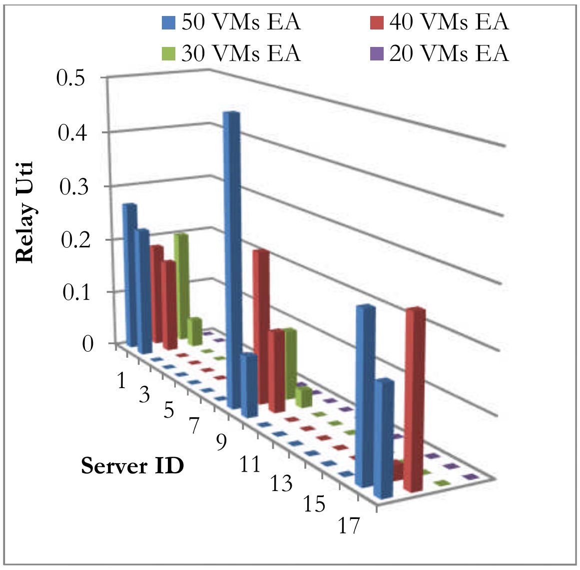

Figure. 3 location and relay utilization of servers selected for routing for the energy aware objective

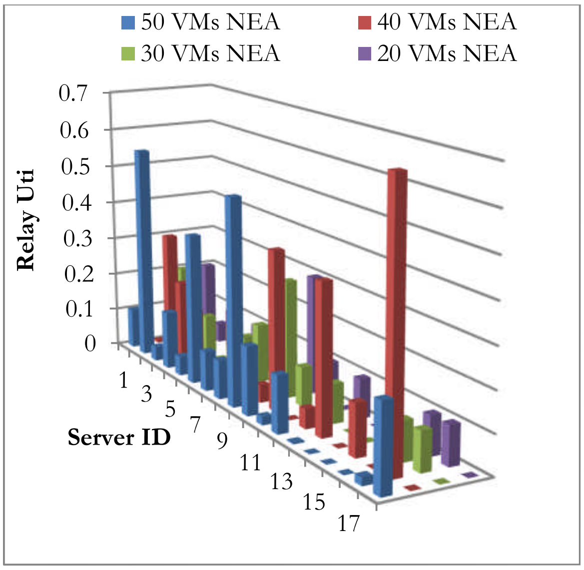

Figure.4 location and relay utilization of servers selected for routing for the non energy aware objective (random placement)

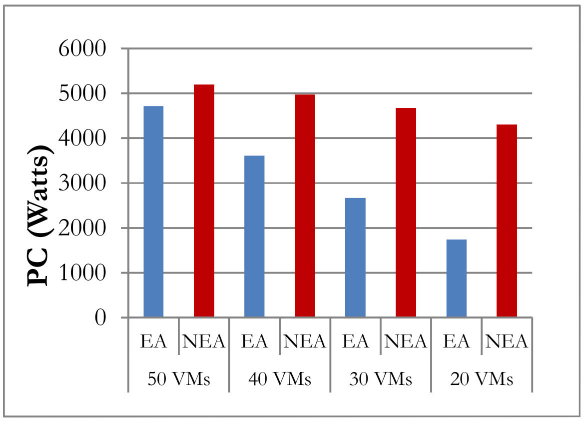

Figure. 5 Total power consumption for energy aware (EA) and non energy aware (NEA) models

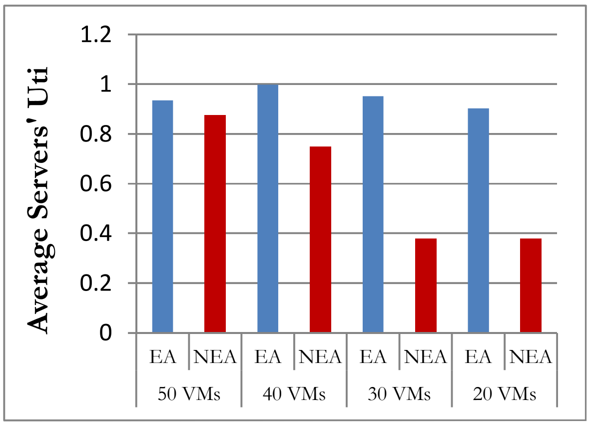

Figure. 6 Average servers' utilization (routing and processing) for energy aware (EA) and non energy aware (NEA) models

## 4. Conclusions

In this paper we have presented a high performance server centric architecture design based on PONs for future cloud data centers. We considered the different merits of PONs in the design of energy efficient, high capacity, low cost, scalable, and highly elastic networking infrastructures that supports the applications and services hosted by modern data centers. We avoided the use of power hungry devices such as switches and tuneable lasers and encouraged the use of low power passive optical backplanes to facilitate intra-rack communication and PONs to facilitate inter-rack communication. We have presented our MILP model results for energy efficient routing and resource provisioning within the PON cell. We optimized the selection of hosting servers, routing paths and relay servers to achieve efficient resource utilization reaching 95% and optimum saving in energy consumption reaching 59%.

## Acknowledgments

The authors would like to acknowledge funding from the Engineering and Physical Sciences Research Council (EPSRC), INTERNET (EP/H040536/1) and STAR (EP/K016873/1).

## References

[1] A. Hammadi and L. Mhamdi, "Review: A survey on architectures and energy efficiency in Data Center Networks," Computer Communication. vol. 40, pp. 1-21, 2014.

[2] X. Dong, T. El-Gorashi, J.M.H. Elmirghani,"IP Over WDM Networks Employing Renewable Energy Sources", IEEE Journal of Lightwave Technology, vol.29, no.1, pp. 3-14, Jan. 2011.

[3] X. Dong, T. El-Gorashi, J.M.H. Elmirghani, "Green IP Over WDM Networks With Data Centers", IEEE Journal of Lightwave Technology, vol.29, no.12, pp.1861-1880, June 2011.

[4] X. Dong, T. El-Gorashi, J.M.H. Elmirghani, "On the Energy Efficiency of Physical Topology Design for IP Over WDM Networks", IEEE Journal of Lightwave Technology, vol.30, no.12, pp.1931-1942, June 2012.

[5] A.Q. Lawey, T. El-Gorashi, J.M.H. Elmirghani, " Distributed Energy Efficient Clouds Over Core Networks", IEEE Journal of Lightwave Technology, vol.32, no.7, pp.1261-1281, April 2014.

[6] N.I Osman, *et al.* "Energy-Efficient Future High-Definition TV", IEEE Journal of Lightwave Technology, vol.32, no.13, pp.2364,2381, July 2014.

[7] A.Q. Lawey, T. El-Gorashi, J.M.H. Elmirghani, " BitTorrent Content Distribution in Optical Networks", IEEE Journal of Lightwave Technology, vol.32, no.21, pp. 4209-4225, Nov 2014.

[8] L. Nonde, T. El-Gorashi, J.M.H. Elmirghani, "Energy Efficient Virtual Network Embedding for Cloud Networks", IEEE Journal of
Lightwave Technology, vol.33, no.9, pp. 1828-1849, May 2015.

[9] A. Hammadi, Mohamed O. I. Musa, T. E. H. El-Gorashi, and J.M.H. Elmirghani, "Resource Provisioning for Cloud PON AWGR-Based Data Center Architecture", 21st European Conference on Network and Optical Communications (NOC), Portugal, 2016.

[10] J. M. H. Elmirghani, Hammadi, A. and El-Gorashi, T.E., "Data Centre Networks", UK patent, 26 November 2014.

[11] A. Hammadi, T. E. H. El-Gorashi, and J.M.H. Elmirghani,"High Performance AWGR PONs in Data Centre Networks," IEEE ICTON, Hungray, 2015.

[12] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture", in Title of Proceeding, ACM
SIGCOMM, Seattle, WA, USA, 2008.

[13] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, "BCube: a high performance, server-centric network architecture for modular data centers", SIGCOMM Computer Communication, vol. 39, pp. 63-74, 2009.

[14] J. Beals IV, N. Bamiedakis, A. Wonfor, R. Penty, I. White, J. DeGroot Jr, et al., "A terabit capacity passive polymer optical backplane based on a novel meshed waveguide architecture," Applied Physics A, vol. 95, pp. 983-988, 2009.

[15] D. Kliazovich, P. Bouvry, and S. U. Khan, "GreenCloud: a packet-level simulator of energy-aware cloud computing data centers," The Journal of Supercomputing, vol. 62, pp. 1263-1283, 2012.