# Towards Realizing the Importance of Placing Fog Computing Facilities at the Central Office of a PON

S.H. Shah Newaz[1], Wida Susanty binti Haji Suhaili[1], Gyu Myoung Lee[3, 4], Mohammad Rakib Uddin[2],
Alaelddin Fuad Yousif Mohammed[3], and Jun Kyun Choi[3]

[1] School of Computing and Informatics, Universiti Teknologi Brunei (UTB), Gadong, Brunei Darussalam.
[2] Dept. of Electrical and Electronic Engineering, Universiti Teknologi Brunei (UTB), Gadong, Brunei Darussalam.
[3]Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea.
[4] School of Computing and Mathematical Sciences, Liverpool John Moores University
Liverpool, United Kingdom.
Email: {shah.newaz, wida.suhaili, rakib.uddin}@utb.edu.bn, G.M.Lee@ljmu.ac.uk, {alaelddin, jkchoi59}@kaist.ac.kr.

*Abstract*—This paper highlights the importance of accommodating Fog computation capability in a Passive Optical Network (PON). Fog computation is referred to as a decentralized computing infrastructure at the network access segment so as to facilitate computation capability closer to end users. PON is seen to be one of the key players in expansion of communication network. PON has paved the way in realizing high speed broadband Internet access to the end users. This paper presents an architectures, namely Fog Co-located Optical Line Terminal (FC-OLT). The FC-OLT architecture is the first effort towards presenting an integrated architecture of Fog computation facilities and PON. This architecture is presented considering Time Division Multiplexing (TDM)-PON (TDM-PON) topology. Along with explaining importance of this architecture in improving QoS/QoE parameters, we point out several important research issues in order to provide direction for future research. Initial performance measurement of FC-OLT presented in this paper reveals that FC-OLT architecture can contribute a great deal of in improving traffic performance.

*Index Terms*—TDM-PON, Fog computing, traffic latency, virtualization.

## I. INTRODUCTION

Passive optical network (PON) technology is considered to be one of the ideal choices for network operators for expanding their footprints in access network segments because PON technology introduces several benefits including high scalability, rich bandwidth capacity, cost-effective services and relatively low energy consumption compared to other access network technologies [1], [2]. Due to wide penetration of mobile devices and stupendous growth of rich Internet applications, PON is considered to be an ideal choice as backhaul for wireless access segment. Among different PON technologies, Time Division Multiplexing (TDM)-PON (TDM-PON) is popularly deployed across different regions of the world.

A TDM-PON comprised of an Optical Line Terminal (OLT), a *1*N* passive star coupler (SC), and several Optical Network Units (ONUs), which are placed at the customer premises. From the OLT, an ONU could be placed 20 *km* or farther away. In a TDM-PON, the OLT, which is the master node, forwards downlink traffic in a broadcast manner. Whereas, all the ONUs, are considered as slave node, can

send uplink traffic during their assigned transmission slots. All the incoming traffic from the customer premises are initially stored at the buffer of an ONU. In turn, the ONU measures the amount of bandwidth it would be requiring for transmitting all the traffic in its buffer and notifies the OLT. The OLT gets all these bandwidth requests from all the subordinate ONUs during a predefined amount of time. After the OLT has knowledge on the amount of bandwidth that each ONU needs, it uses a bandwidth allocation algorithm to measure uplink transmission slots for the ONUs. On the other hand, an ONU manages priority based scheduling to meet the Quality of Service (QoS) requirement of different types traffic in its buffer [2], [3].

In case of cloud computing, which facilitates computation and storage facilities for the clients, latency is also an important factor. The major shortcoming of cloud is that cloud might fail to meet required level of latency for the users. To minimize latency between two end points (e.g. a cloud server and end user), it is increasingly important to manage their data flows efficiently. Indeed, distance between a cloud server and end user can play a crucial role in determining how much delay a particular flow might experience. In order to meet delay requirements of delay sensitive applications that need support of cloud computation facilities, researchers have come up with the concept of Fog computing, which puts computational load at edge of network, thus reducing the distance between a computing facilities and a client.

Servers deployed for Fog computing usually located at the network *Tier 2*, whereas cloud servers are located at the *Tier 3* [5]. In 2012, Cisco introduced the concept of Fog computing [4]. Fog computing offers several advantages over the conventional cloud computing, including, a) short end-to-end latency, b) low jitter, and c) small amount of packet loss.

Currently, the main research streams in the Fog computing are: (i) scalable resource allocation, (ii) scalable reliability management, (iii) dynamic on demand security management, (iv) scalable software and management infrastructures, and (v) security and trust management [6].

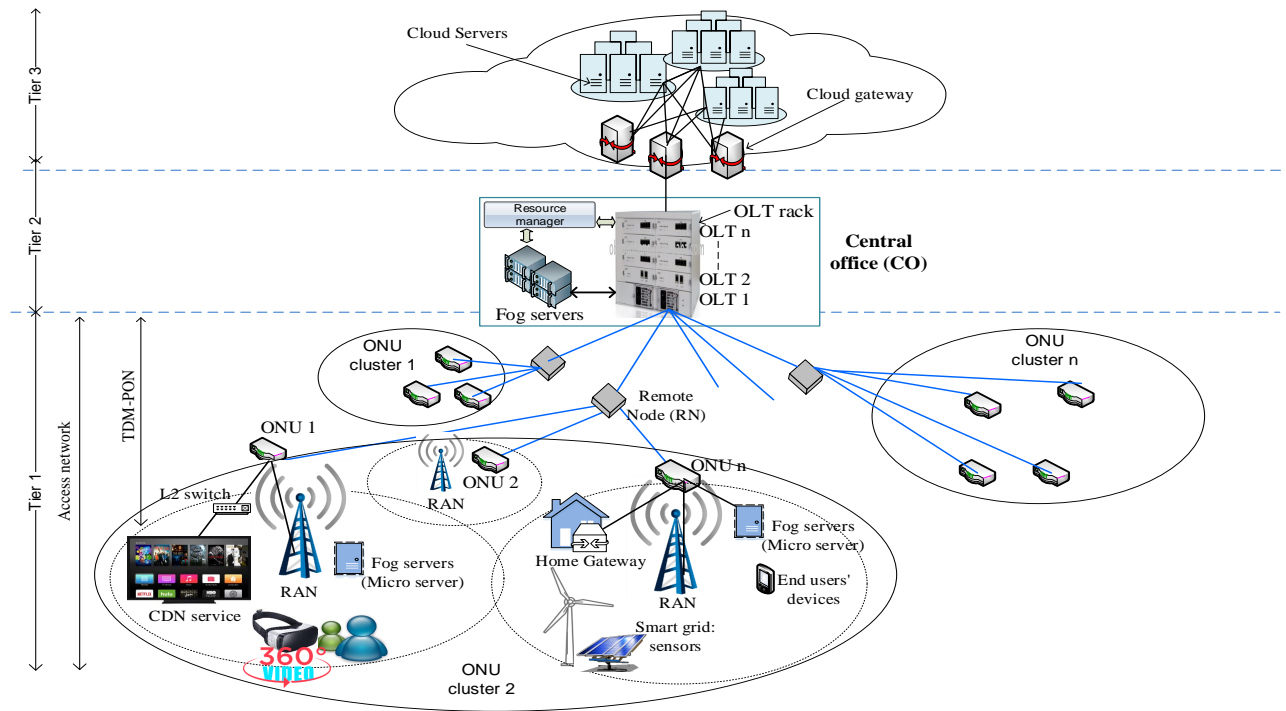In this paper, we present an integrated architecture of Fog

Fig. 1. Fog Co-located OLT architecture: serving numerous applications of customer premises.

servers and OLT of a TDM-PON. Need to highlight that PON has been widely adopted throughput the world and its footprint in expanding Internet coverage in access segment is still stupendously growing. An ONU in a TDM-PON usually connects on an average 16 homes. Therefore, placing Fog computational facilities close to a OLT can lead to have several advantages, including local content caching, number of packet loss reduction, proving computational support to the end users (code offloading) and meeting QoS requirements of end users' applications.

In particular, in this paper, we present an architectures, namely Fog Co-located OLT (FC-OLT). Performance measurement based on OPNET simulator reveals that FC-OLT architecture can contribute a great deal of in improving traffic performance.

The rest of the paper is organized as follows. Section II presents related work. Section III introduces FC-OLT architecture and explains how this architecture can contribute in improving traffic performance. Section IV shows performance evaluation results of our solution. Finally, Section V concludes this paper.

## II. RELATED WORK

Introducing Fog computing servers at the network edge brings several advantages. To get full potential of Fog computation, there are several research efforts can be noticed. One of the important focuses is Fog server placement in order to reduce Capital Expenditure (CAPEX) and Operational Expenditure (OPEX), while meeting QoS of applications running at

the end users' terminals. In [8], authors argue that a network operator needs to understand the number of computational and storage load for each of the Fog servers. That is, before deploying Fog servers co-located with RANs in order to support mobile users, a network operator needs to anticipate the maximum computational and storage load during the peak hours of a day so as to meet the demand of end users.

Automatic resource scaling in Fog commutation is crucially important. A novel automated scaling approach is presented in [9], [10]. Authors came up with two scaling framework, namely SmartScale and WebScale in order to facilitate automating scaling in horizontal and vertical so as to optimize resource usage while minimizing cost.

Authors in [11] present a gateway based Fog computing architecture. In their architecture a gateway and micro server are co-located facilitating a platform in order to provide Internet of Things services and applications along with virtualization of wireless sensor networks.

## III. INTEGRATING FOG COMPUTING IN PON AND ITS FEATURES

This section first presents Fog Co-located OLT (FC-OLT) architecture. Then we discuss several important benefits that FC-OLT can offer in improving QoS of different applications running at users' terminal.

### A. Fog Co-located OLT (FC-OLT) Architecture:

In FC-OLT architecture, Fog servers are co-located with the OLT, which is situated at the central office (CO) of a TDM-PON system, as shown in Fig. 1. In CO, usually several OLTs

are placed within one or more racks. Each of the OLTs is in charge of several ONUs. In our architecture, a Fog Server can be connected with one or more OLTs. At a given time, the number of active Fog server mainly depends on demand of access segment. The *Resource manager* located at the CO is in-charge of allocated resource in the Fog servers located in the CO. These Fog servers accommodate different applications aside from storage functionality.

As we can notice from Fig. 1 in the customer premises, under each ONU cluster there could be several applications running at the end users devices including smart traffic control, smart grid and Content Delivery Network (CDN) service [12]. Next, we explain an operational mechanism of *Resource manager* and an OLT in this proposed FC-OLT architecture.

*1) Resource Manager:* It is in charge managing computational and storage resources both in Fog servers and in the access segments (if there is any). *Resource manager* is also aware of task deadline when any task is assigned at one of the co-located Fog servers. To meet task completion deadline, it takes into account both task processing delay at the Fog server and the transmission delay from the OLT (CO) to a particular end user terminal.

*Resource manager* also facilitates resource virtulization and orchestration in the access segment.

*Resource manager* manages synchronization with the cloud server located in *Tier 3* and the Fog servers located in access segment. It also distributes task among these distributed resources taking into account different QoS factors (for example, in cases of video streaming service it should take into account startup delay, throughput and jitter).

In case when the For servers in FC-OLT architecture becomes part of CDN service provision, the *Resource manager* needs to decide on content caching location (content storage location could be in Fog servers or cloud server depending on QoS requirement).

*2) Optical Line Terminal (OLT):* Aside from performing all the functionalities that a conventional OLT does, in FC-OLT, an OLT needs to strongly collaborate with the *Resource manager* in order to meet QoS. For example, if the *Resource manager* invokes the OLT to allocate more bandwidth for a particular flow so as to compensate delay when a Fog server misses deadline, the OLT needs to reschedule traffic transmission.

### B. FC-OLT Features

Here we list the main advantages that the proposed architecture offers:

- In FC-OLT, in case of CDN content caching in Fog servers, it will be possible to improve QoS performance of traffic. Additionally, *Resource manager* can manage selection of content for caching and content location dynamically. For example, *Resource manager* can place a content either at the Fog servers in the CO or any micro server (Fog servers) located at the customer premises based on the content popularity.

Putting Fog servers co-located with the OLT facilitate *Resource manager* to have easy and almost real-time control over the resources reside in different ONU clusters.

Real time control and service provisioning to vary latency sensitive applications running at the customer premises. For example, Fog servers at the CO can serve smart traffic control and smart grid applications at the customer premises.

In FC-OLT, a Fog server could be part of virtual content storage location for any of the home gateway located in the customer premises.

- In case of mobile video streaming, *Resource manager* in FC-OLT architecture can easily relocate the video content whenever required. This will leverage QoS performance in such mobile video streaming scenario. Synchronization and dynamic flow adaption are increasingly important for reliable multi-stream transmission [15]. The proposed FC-OLT can easily handle multi-steam transmission scenario.

For 360° Virtual Reality (VR) service latency is a crucially important issue to ensure Quality of Experience (QoE). For example, Samsung Gear VR requires to have Motion to Photon should be Less than 20 ms [16]. In this proposed architecture, the *Resource manager* should be able to dynamically initiate video streaming so as to meet Motion to Photon requirement.
- Resource virtualization and easy orchestration can be actualized easily when our FC-OLT is in place. For example, it will allow easy multiple screen synchronization and real time interactive media control at the end users' display.
- Easy communication protocol conversion.
- Integrating both Fog servers with the OLT will lead to have more control maintaining QoS of applications.
- Dynamic packetization (variable FEC and AL-FEC) is possible to be implemented. Considering network condition in wireless access segment, the length of FEC code can be varied so as to let the end terminal recover any erroneous information from the received frame while not wasting bandwidth significantly.
- Dynamic load balancing taking account the traffic condition in the access segment [13].
- FC-OLT can easily facilitate Layer 3 location update and handover management for the mobile terminals in the wireless access segment. Co-located Fog computational facility can house virtualized Local Mobility Anchor (LMA) or Mobile Access Gateway (MAG) functionality [13]. Earlier research presented in [13] demonstrates that such an architecture can lead to provide several advantages including signaling cost minimization, reduction of start up latency, and reduction of packet loss during

handover of mobile terminals.
- In case when Software Defined Networking (SDN) based traffic flow in access segment is managed [17], in FC-OLT, a Fog server can also accommodate SDN controller functionalities.

## IV. RESULTS AND DISCUSSION

To highlight importance of the presence of Fog computing servers co-located with the OLTs at a CO, we have evaluate FC-OLT architecture in our outstanding TDM-PON OPNET based simulation model which has been earlier used in many of our research efforts (e.g. [19]). We conducted simulation under two scenarios: (i) FC-OLT and (ii) Conventional Cloud Service Architecture (CCSA), in which all the cloud services are processed at the *Tier 3*.

Here, we present the assumptions for performance analysis. For CCSA architecture, in our simulation, we consider that the Round Trip Time (RTT) between a cloud server and the OLT is 100 *ms*. For both CCSA and FC-OLT, we consider that the OLTs uses dynamic bandwidth allocation Interleaved Polling with Adaptive Cycle Time (IPACT) scheme. Furthermore, we assume that the average distance between the CO and an ONU is 20 *Km*. The link speed between an OLT and ONU is 1 *Gbps* and each of the ONUs serves 16 users, as considered in [2].

Figure 2 demonstrates upload response time and through results for both FC-OLT and CCSA under wide range of ONUs. Looking at this Fig. 2 (a) we can comprehend that, in case of FC-OLT, the upload response time results are noticeably smaller than that of the results under CCSA regardless the number of ONUs connected with an OLT. Throughput performance of both of them is depicted at Fig. 2 (b). Here, again, we can notice that FC-OLT outperforms in terms of throughput compared to CCSA.

This is an initial study in an area where little has previously been done considering the existence of Fog computational facilities along with an OLT rack at a CO. Here, we highlight several practical concerns toward implementing FC-OLT.

- OLTs and *Resource manager* need to have strong collaboration to meet QoS requirements of different application running at the customer premises.
- For very latency sensitive applications (e.g. smart grid communication, road traffic control, VR video streaming), the OLT needs to adjust both downlink and uplink bandwidth allocation policy dynamically.
- Sleep mode has been widely studied in the PON research area so as to maximize energy efficiency of PON equipment (e.g. ONUs) [2], [18], [19]. The OLT needs to set sleep interval lengths for the ONUs taking into account latency sensitive applications served by the Fog servers. If in case, sleep to active state transition is relatively long (e.g. 5 *ms*), the OLT should not allow an ONU to sleep in presence of latency sensitive traffic flows.
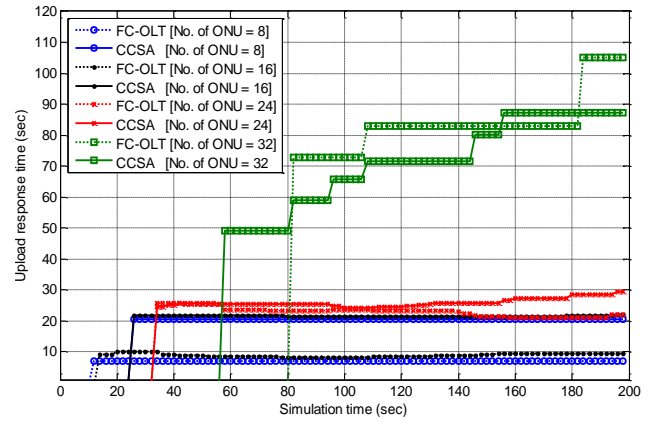


(a) Upload response time
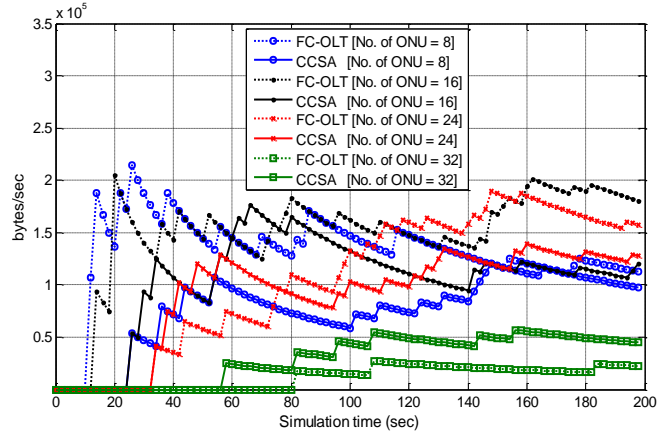


(b) Throughput

Fig. 2. Upload response time and throughput comparison.

When very latency sensitive traffic flow is available for an ONU, the OLT might invoke the ONU to use doze mode (in this energy saving mode of an ONU, an ONU turns off the transmitter module, while powering up receiver module [2], [18]), thereby allowing the OLT to forward downlink traffic whenever it has any traffic for that ONU.

- Seamless mobility management is increasingly important for many applications (e.g. mobile video streaming). The *Resource manager* needs to take into account both inter OLT and intra OLT mobility of end users' device while forwarding downlink traffic.
- As mentioned earlier FC-OLT should be able to facilitate distributed content caching. Dynamically caching location selection in access segment taking into account end users' device mobility is challenging undoubtedly. The *Resource manager* needs to be aware of the attached end users' devices with an ONU at a given time. Additionally, mobility prediction of mobile users could be a useful technique in order to relocate favourite contents beforehand. This will result in improving performance of applications with strict QoS requirements. Furthermore, bicasting approach can be adopted in FC-OLT both during

inter and intra OLT handoff in order to minimize packet loss [20].

- In case when computational load is distributed across different computational facilities (e.g. micro servers, cloud servers in *Tier 3*), the *Resource manager* needs to take into account traffic forwarding delay among different computational facilities before setting task completion deadlines.

## V. CONCLUSION

This paper proposes FC-OLT architecture. We have provided detailed discussion on how the proposed architecture can contribute in improving QoS and how it facilitates different distributed resources in the customer premises. In fact, for the latency sensitive applications having Fog server integration with the OLT, which is located at the CO, is increasingly important. To the best of our knowledge this is the first effort towards presenting an architecture that combines Fog computation capability with PON. In our future research we will present detailed performance results in order to demonstrate potential of our proposed architecture.

## REFERENCES

[1] S. H. S. Newaz *et al.,* "Energy efficient and latency aware TDM-PON for local customer internetworking". In IFIP/IEEE International Symposium on Integrated Network Management (IM), pp. 1184-1189, 2015.

[2] S. H. S. Newaz *et al.,* "Adaptive delay-aware energy efficient TDM-PON", In Computer Networks, vol. 57, pp. 1577-1596, 2013.

[3] IEEE Std. 802.3ah-2004, "Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications," Institute Computer Society, 7 Sept. 2004.

[4] F. Bonomi *et al.*, "Fog Computing and Its Role in the Internet of Things," *in Proceedings of ACM MCC*, 2012, pp. 13-16.

[5] S. Sarkar; S. Chatterjee; S. Misra, "Assessment of the Suitability of Fog Computing in the Context of Internet of Things," *in IEEE Transactions on Cloud Computing*, no.99, pp. 1-1, 2015.

[6] OpenFog, 2016. OpenFog Architecture Overview. Available at: https://www.openfogconsortium.org/white-paper-reference-architecture/white-paper-download-open-fog-reference-architecture

[7] Hung, S.C. *et al.,* "Architecture harmonization between cloud radio access networks and fog networks". *IEEE Access*, 3, pp. 3019-3034, 2016.

[8] Luan, T.H. *et al.,* "Fog Computing: Focusing on Mobile Users at the Edge". eprint arXiv:1502.01815, pp.1-11. Available at: http://arxiv.org/abs/1502.01815.

[9] Dutta, S. *et al.,* "SmartScale: Automatic application scaling in enterprise clouds", *In proceedings IEEE 5th International Conference on Cloud Computing*, CLOUD 2012, pp. 221-228, 2012.

[10] Lin, C.C. *et al.,* " Automatic resource scaling based on application service requirements," *In Proceedings - 2012 IEEE 5th International Conference on Cloud Computing*, CLOUD 2012, pp. 941-942, 2012.

[11] W. Lee, *et al.,* "A gateway based fog computing architecture for wireless sensors and actuator networks," *18th International Conference on Advanced Communication Technology (ICACT)*, pp. 210-213, 2016.

[12] Liang Chen *et al.,* "Thunder crystal: a novel crowdsourcing-based content distribution platform", *In Proceedings of the 25th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, 2015.

[13] S. H. S. Newaz *et al.,* "An approach of multipath transmission from intelligent OLT in optical-wireless converged networks", Digest of the 9th International Conference on Optical Internet (COIN 2010), pp. 1-3, 2010.

[14] S. H. S. Newaz *et al.,* "Efficient IP Mobility Management for Green Optical and Wireless Converged Access Networks," *IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1-5, 2012.

[15] Chung-Ming Huang *et al.,* "Synchronization and flow adaptation schemes for reliable multiple-stream transmission in multimedia presentations", *Journal of Systems and Software archive*, vol. 56 Issue 2, pp 133-151, Mar. 2001.

[16] Samsung Gear VR Innovator Edition (2016). [Online] Available:http://xinreality.com/wiki/Samsung_Gear_VR_Innovator_Edition.

[17] W. Xia *et al.,* "A Survey on Software-Defined Networking," *in IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 27-51, Firstquarter 2015.

[18] ITU-T Recommendation G Suppl. 45, Supplement on GPON, Power Conservation, May 2009.

[19] A. F. Y. Mohammed *et al.,* "Early wake-up decision algorithm for ONUs in TDM-PONs with sleep mode", in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 8, no. 5, pp. 308-319, 2016.

[20] M. I. Sanchez *et al.,* " On providing mobility management in WOBANs: integration with PMIPv6 and MIH," *in IEEE Communications Magazine*, vol. 51, no. 10, pp. 172-181, Oct. 2013.