QAMO-SDN: QoS Aware Multipath TCP for Software Defined Optical Networks

Sana Tariq and Mostafa Bassiouni
Department of Elec. Eng. & Computer Science
University of Central Florida, Orlando, FL USA
sana.tariq@knights.ucf.edu, bassi@cs.ucf.edu

Abstract- Cloud based datacenters will be most suitable candidates for future software defined networking. The QoS requirements for shared data centers, hosting diverse applications could be successfully achieved through SDN architecture. This paper provides an extension of our previously proposed scheme QAMO that was aimed at achieving tangible QoS in datacenters through controlling bandwidth reservation in Multipath TCP and OBS layer while maintaining throughput efficiency. However, QAMO was designed for traditional networks and did not have the capability to adapt to current network status as expected from future software defined networks. The paper presents an enhanced algorithm called QAMO-SDN that introduces a controller layer in previously proposed architecture and achieves adaptive QoS differentiation based on current network feedback. QAMO-SDN inherits the architecture of QAMO, using Multipath TCP over OBS networks. We evaluate the performance of QAMO-SDN under different network loads and topologies using realistic data center traffic models and present the results of our detailed simulation tests.

***Index Terms* - MPTCP, all-optical network, data center, TCP, OBS, QoS, SDN**

I. INTRODUCTION

Many internet applications today are powered by data centers equipped with hundreds of thousands of servers. The concept of shared datacenters also became popular with the widespread adaptation of cloud. There is a growing interest in introducing QoS (Quality-of-Service) differentiation in datacenters, motivated by the need to improve the quality of service for time sensitive datacenter applications and to provide clients with a range of service-quality levels at different prices. There is also a growing trend towards software defined networks in datacenters and QoS schemes should adapt to SDN based cloud architectures. Software defined networks could be well understood by a simple analogy of sending a package through a courier who sets off his way in the network to deliver it. Traditionally he would ask different people and change his route multiple times to find an optimal path. With a Software defined network, assume that the courier has a GPS system with an up to date data of all the routes, traffic conditions, packet size and its requirements to find the best route for it dynamically before it is launched into the network. In the traditional network the routers contain the rules and logic for controlling the flow and modifications of packets. In traditional network there no centrally controlled mechanism to route the traffic. Software defined networks decouples the control plane from the data plane so the packet traverse the network with a pre-defined knowledge of the route it will take and the control of the traffic lies within the software defined network controller. Achieving QoS through current network feedback as expected in software defined networks will improve the performance and efficiency of QoS schemes.

The type of applications hosted by datacenters are diverse in nature ranging from back-end services such as search indexing, data replication, MapReduce jobs to front end services triggered by clients such as web search, online gaming and live video streaming [1]. The background traffic contains longer flows and is throughput sensitive while the interactive front end traffic is composed of shorter messages and is delay sensitive. The traffic belonging to the same class can also have differences in relative priority levels and performance objectives [2].

In this paper, we employ MPTCP over OBS for datacenter networks for efficiency and robustness as was done in our previous work [3] and present and evaluate a QoS provisioning algorithm in software defined networks called QAMO-SDN, 'QoS aware MPTCP over software defined optical network'. To our knowledge, this is the first research report that provides QoS provisioning algorithm for service differentiation using MPTCP over OBS in software defined datacenter.

The rest of the paper is organized as follows. In section II, we review previous work. In section III, we describe our networking model that uses MPTCP protocol over an optical burst switching network for data centers. In section IV, we present 'QoS Aware Multipath TCP for Software Defined Optical Networks', QAMO-SDN scheme. Simulation details are discussed in section V and the performance analysis and experimental results are given in Section VI. We conclude the paper in Section VII.

Shared data center applications constitute a complex mix of workloads from multiple organizations. Some workloads require small predictable latency while others require large sustained throughput. Such shared data-centers are expected

978-1-4799-6390-4/15/$31.00

to provide QoS to client's individual flows and with the widespread popularity of software defined networking the QoS techniques should also to adapt to SDN architecture of future networks.

II. PREVIOUS WORK

Cloud services are expanding and organizations are shrinking their datacenters to virtualization technologies in order to take advantage of the predictability, continuity, and quality of service. Future data center consumers will require quality of service QoS as a fundamental feature. Software defined networks have received significant attention in industry and research community in recent years [4-9]. There have been some recent research studies on traffic modeling, network resource management and QoS provisioning in data centers [1, 10, 11]. Ranjan, et. al., studied the problem of QoS guarantees in data-center environments in [11]. However, this work is not suitable for highly loaded shared data-centers with computationally intensive applications due to the two sided nature of communication. Song Ying et al. in [12] proposed a resource scheduling scheme which automatically provides on-demand capacities to the hosted services, preferentially ensuring performance of some critical services while degrading others when resource competition arises. Hong et al. in [13] proposed a flow scheduling protocol called Preemptive Distributed Quick (PDQ), designed to complete flows quickly by emulating a shortest job first algorithm and giving priority to the short flows. Similarly authors in [14] propose taxonomy to categorize existing works based on three main techniques, reducing queue length, prioritizing mice flows, and exploiting multi-path. Zats et al. in [15] proposed DeTail, which designed a cross-layer network stack aiming to improve the tail of completion time for delay-sensitive flows. Wilson et al. [16] presented a deadline-aware control protocol, named D3, which controlled the transmission rate of network flows according to their deadline requirements. D3 gave priority to mice flows and improved the transmission capacity of data center networks.
Previously, research studies on QoS provisioning in data centers did not employ optical networks nor did they use multi-access transport protocols such as MPTCP. Extending our previous work on MPTCP over OBS in datacenters [3], and QoS provision scheme QAMO, allows us to create a network model and QoS provisioning scheme for software defined datacenters called QAMO-SDN.
There is rich research on QoS schemes in optical burst switching for wide area networks [17-19]. OBS has been considered as the best compromise between optical circuit switching (OCS)and optical packet switching (OPS) due to its granularity and bandwidth flexibility, and would be suitable for data centers eventually as optical switching technology gets mature [20]. TCP is the most dominant transport layer protocol in internet and TCP over OBS has been extensively studied [21-23]. Multipath-TCP (MPTCP) has been shown to provide significant improvement in throughput and reliability in electronic packet switched networks in data centers [24, 25]. However, MPTCP has not been studied in the context of OBS networks before.

III. NETWORK MODEL

Since SDN separates the control plane and data-forwarding plane, the entity that implements the control-plane functionalities is called the SDN controller. Software defined network has SDN capable devices hence software defined optical network will have SDN enabled optical cross connects that can communicate with upper layers[4, 26]. With an SDN architecture the controller layer has a lower level network view that enables the QoS schemes to perform prioritization of flows based on actual bandwidth on the links and network state. Figure 1 below shows the high level diagram of software defined network architecture.

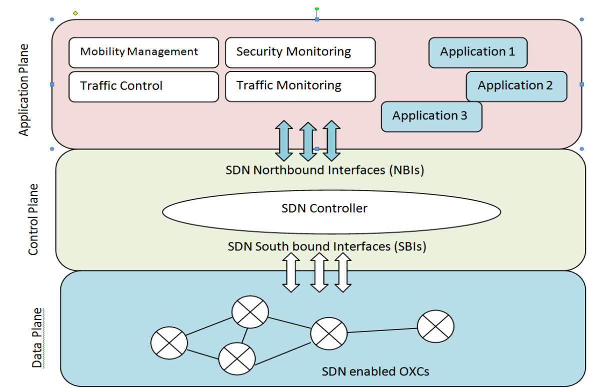

Figure 1: High level SDN architecture

With the popularity of new data center topologies such as Fat Tree and VL2 and the multitude of available network paths, it becomes natural to switch to multi path transport protocol such as MPTCP to seek performance gains. MPTCP provides significant improvement in bandwidth, throughput and fairness. We have used MPTCP over OBS in our proposed network architecture. In an OBS network, the control information is sent over a reserved optical channel, called the control channel, ahead of the data burst in order to reserve the wavelengths across all OXCs (Optical cross connects).In our SDN architecture we assume that our optical cross connects will be SDN enabled and will have the functionality to communicate the available wavelengths with upper layers [27]. The wavelength reservation protocol plays a crucial role in the burst transmission and we have used just-in-time (JIT) [28] for its simplicity. The necessary hardware level modifications of optical switches for supporting OBS in data centers have been discussed in [29], and will not be repeated in this paper.

IV. QOS AWARE MPTCP OVER OBS ALGORITHM

Our proposed algorithm QAMO-SDN combines the multiple paths of MPTCP and resource reservation in OBS to develop an adaptive and efficient QoS-aware mechanism. Data centers handle a diverse range of traffic generated from different applications. The traffic generated from real time applications e.g., web search, retail advertising, and recommendation systems consists of shorter flows and requires faster response. These shorter flows (foreground traffic) are coupled with bandwidth intensive longer flow (background traffic) carrying out bulk transfers. The bottleneck created by heavy background traffic impacts the performance of latency sensitive foreground traffic. It is extremely important to provide a preferential treatment to time sensitive shorter flows to achieve an expected performance for data center applications. QoS technologies should be able to prioritize traffic belonging to more critical applications. Our proposed algorithm provides priority to latency-sensitive flows at two levels, i) MPTCP path selection stage and ii) OBS wavelength reservation stage. We propose that larger bandwidth be dynamically allocated to high priority flows, in order to minimize latency and reduce their drop probability. Datacenter networks are continuously changing and the concept of software defined networking is becoming increasingly popular. QoS algorithm should adapt to current network and dynamically change routing decisions that achieves service differentiation for current network state. QAMO-SDN algorithm just does that.

Let W be the maximum number of wavelengths per fiber, and K be the number of paths that exist between a given souce-destination pair. We will introduce a new term, the *priority factor P* for a burst priority defined as the ratio of P_{curr} (priority level of the current burst) to P_{max} (maximum priority levels) i.e., $P = P_{curr}/P_{max}$. Priorities of individual bursts are represented in ascending order as P_1, P_2, $P_{3...}$ P_{max} while P_{max} is the highest priority level in the bursts. As discussed before, the number of allocated paths for the burst of a particular priority level as follows.

$$\text{max\_paths} = \lceil K \times P \rceil \qquad (1)$$

We also define a new vector $Path_{i,j}$ which is a collection of all paths that exist between nodes *i* and *j*. The number of paths this vector must store can be limited based on set value of *K* to reduce overhead. Another matrix is introduced in the algorithm, *L*, link state matrix. Each element $L_{i,j}$ of the matrix shows the state of the link between the nodes *i* and *j* as below:

$$L_{i,j} = \textit{number of available wavelengths between nodes i and j} / W \qquad (2)$$

$L_{i,j}$ is initialized to 1 as all wavelengths are available and as the networks becomes congested, the matrix gets updated as shown in equation (2). We then sort the vector of paths in descending order of available bandwidths along the path. So the path having higher availability will be put on top of the path having lower number of paths. Since the number of paths in path vector can be limited, only the shortest paths will be chosen and then arranged accordingly. Lightpath creation is done only on the subset of paths from this vector. This subset is chosen from the top, so the path having higher wavelength availability will be preferred over others. At path allocation stage a larger number of paths is allocated for a high priority burst thus reducing its latency. For example, if P_{curr}=P_{max}, then P = 1. This will result in k_{curr} = K paths whereas if $P_{curr} = 0.5*P_{max}$, then $P = 0.5$ and the number of allocated paths is reduced to half the set of *K paths*. This will give the low priority burst, half the number of paths. We now define the size of the wavelength search space controlled by the following equation.

$$\textit{Wavelength search size} = \lceil W \times P \rceil \qquad (3)$$

At wavelength reservation stage in OBS, equation 3 allocates a larger subset of wavelength search space for a burst with higher priority level thereby allowing it a greater chance to get through and reduce its blocking probability.

After reserving all the lightpaths for current source/destination pair, matrix L is updated according to equation (2) and made available to be used during other reservations.

QAMO (QoS Aware MPTCP over OBS) Algorithm

Input:

```
        P = P_cur/P_max
        K = maximum number of paths
        W = maximum number of wavelengths
        w_cur = current wavelength reserved for current burst
        N_k = vector of all nodes on path k
        Paths_i,j = vector of all paths between node i and node j
        burst_cur = current burst
        L = link state matrix
        L_i,j = state of link between node i and node j
        N = vector of all nodes in the network
```

Algorithm:

```
if L is not initialized
  Initialize matrix L: set L_i,j = 1
arrange_paths(i, j)
max_paths = ⌈K × P⌉
for each k in max_paths:
  path_curr = Paths_i,j[k]
  lightpath(path_curr)
update_link_state_matrix()
```

function lightpath(path k)

```
   Initialize w_cur
   for each n in N_k:
     if n = N_k [length(N_k) – 1]             // destination node
       break;
     if n = N_k                               // source node
       for each w in ⌈W × P⌉:
         if w is free
           reserve w for burst_cur at n
           w_cur = w
           break;
     else
       if w_cur is free at n
         reserve w_cur for burst_cur at n
         continue;
       for each w in ⌈W × P⌉:
         if w is free
           reserve w for burst_cur at node n
           w_cur = w
           break;

     if no free wavelength at n
```

```
      return (error);                    // search failed at node n
  return(success);

  function arrange_paths(node i, node j)
    sort all paths in Paths_{i, j} in descending order of average
bandwidth availability on all links along the path

  function update_link_state_matrix()
    for each node i in N
      for each node j in N
        if i = j
          continue;
        L_{i, j} = number of available wavelengths between i and j / W
```

In the above algorithm, the priority factor P is used to adjust the number of allocated paths for concurrent transmission and the size of the wavelength search space based on the priority level of the burst. The chosen paths will always have higher availability compared to others. For high priority bursts, more concurrent MPTCP paths result in larger bandwidth, and more OBS network wavelengths reduce dropping probability. The parameter P_{max} can be flexible to accommodate changes in network statistics over time as bursts of different priority levels are encountered.

Figure 2 shows the cross layer design on QAMO-SDN algorithm. We assume that QAMO-SDN algorithm has access to available information about QoS requirements of different bursts and network conditions to process them correctly.

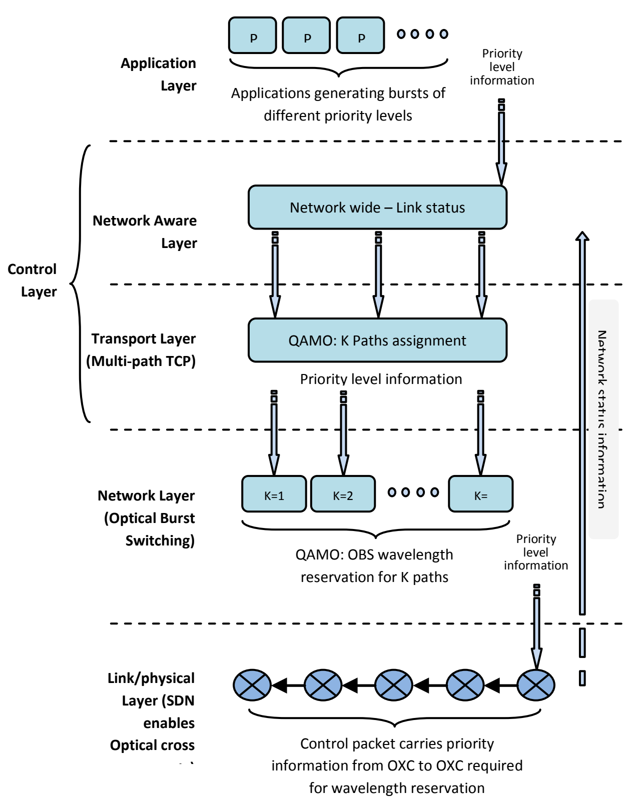

Figure 2: QAMO-SDN's cross-layer design: Changes to the Protocol stack and the burst priority level information flow.

The Controller layer receives feedback from lower layers and establishes an inner view of underlying network topology and stat in terms of link/node congestion. This layer provides feed back to QAMO-SDN layer to calculate the best path for new burst based on its priority level and current situation of wavelengths at OXS along various possible light paths. We have assumed that priority level information will flow from application to MPTCP layer. This capability may be implemented using a specific interface such as the Implicit Packet Meta Header (IPMH) promoted in [30]. It is possible to assign priority levels for different flows in MPTCP at IPMH interface [31, 32]. Because of IPMH interface, it is also possible to gather priority information for each type of flow at a particular end host. This information can be passed on to the OBS network during burst segmentation process from MPTCP layer. At OBS network, the current burst priority P_{curr}, *or the ratio* $P = P_{curr}/P_{max}$, can be easily passed from one SDN capable OXC to the next and upper layers via the control packet and does not demand any significant resources in the OXC's. Implementing the reduced (adjustable) search as in the case of QAMO-SDN, to find a free wavelength requires minor modification to the standard JIT channel allocation scheme. The adjustable search in a smaller space of $\lceil W \times P \rceil$ for wavelengths actually leads to a smaller average search time.The QAMO-SDN scheme has been extensively tested on the simulation testbed using data center network topologies FatTree and BCube and is shown to provide tangible QoS differentiation without negatively impacting the overall throughput of the system. It is also observed that QAMO-SDN utlises available capacity better than basic QAMO scheme due to SDN architecture.

V. SIMULATION DETAILS

The simulation testbed has been developed using C++. A source-destination pair amongst host nodes is randomly chosen for each originated burst. For TCP, to establish the static lightpath, simulation calculates the shortest path between these nodes using Dijkstra's algorithm. In case of MPTCP, it uses K shortest paths algorithm (derived from Dijkstra's algorithm) to find K paths between the source-destination pair. The wavelength assignment heuristic is first-fit as done in [33, 34]. Recent research studies on traffic characteristics of data centers have shown that the traffic in data centers follows the lognormal distribution with ON-OFF pattern [10, 35]. The lognormal distribution is also considered to be the most fitted distribution for modeling various categories of internet traffic including TCP [36]. We have used lognormal arrival with an ON-OFF behavior in our simulation. The network nodes are assumed to be equipped with wavelength converters. We assume that MPTCP is running at end hosts. Based on the priority of the burst, K control packets originate from the source node to establish K lightpaths. Each control packet acquires an initial free wavelength at the source node, then travels to the destination node and reserves wavelengths following QAMO algorithm. If at any node, the same wavelength as the one reserved on the previous node is not available then it tries wavelength conversion. The process continues until the control packet either reaches the destination node or gets blocked due to the unavailability of free wavelength at any hop along the path.

Thus, number of lightpaths established = K – number of control packets blocked. The source node waits for a predetermined time depending on the hop distance to the destination called offset time before transmitting the optical burst message. The traffic used in our simulation is uniformly distributed, i.e., any host node can be a source or a destination [33, 37].

The simulation clock is divided into time units, where each simulation time unit corresponds to 1 microsecond (μs). Each node has a control packet processing time of 20 microseconds and a cut through time of 1 microsecond as proposed for OBS networks in data centers [38]. Each node can have a certain maximum number W of allowed wavelengths. Arrival rate/tu denotes the average arrival rate of the lognormal ON-OFF traffic.In data center environment a complex mix of short and long flows is generated. The shorter flows are usually latency-critical and represent the largest proportion of flows in data centers [10]. The medium sized and longer flows constitute background traffic and may belong to different priority levels [39]. To represent these scenarios of data center mixed traffic, we have used variable burst sizes in different ranges with uniform distribution within each range [39].

Short burst sizes: S_{min}=5 Kbits to S_{max}= 20 KB

Medium burst sizes: S_{min}=200 Kbits to S_{max}= 1 MB

Long burst sizes: S_{min}=20 Mbits to S_{max}= 100 Mbits.

Our traffic model is based on the findings on data center traffic characteristics in [1, 10, 35, 39]. To model our traffic we assume dynamically changing traffic with an average of 70-80% of bursts generated in short burst range belonging to latency sensitive applications, 10-15% in medium burst sizes while 5-10% of bursts belongs to large burst size range. In order to assign the priorities we use dynamically changing priority levels and relative percentages of various priority classes with an average of 95% short burst messages having the randomly assigned priorities from the highest priority range [P5-P6]; the remaining 5% can have any priority level. Similarly, 95% of medium and large burst sizes are randomly assigned priorities from sets [P3 – P4] and [P1 – P2] respectively. The remaining 5% from these ranges are assigned random priorities from set [P1 – P6].

VI. RESULTS AND DISCUSSION

The topologies used in our simulation tests are FatTree with 36 nodes and BCube with 24 nodes as shown in Figure 3 and Figure 4.

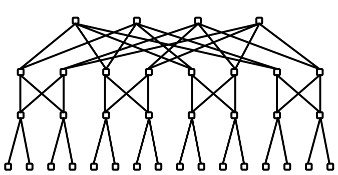

Figure 3. FatTree topology used for simulation

In FatTree topology the root level nodes are called high level aggregators (HLAs), the next layer of nodes are medium level aggregators (MLAs). The longest lightpath in the 36-node FatTree network has the diameter of 6 hops. There are 16 hosts as shown in Figure 3 in the bottom layer..

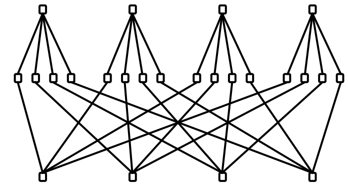

Figure 4. BCube topology used for simulation

The BCube network has 16 relaying hosts in the middle layer. The network diameter in the 24-node BCube network is 4 hops. All the figures in this section are tested following lognormal distribution. Because of the ON-OFF pattern of traffic the average arrival rate is smaller than the arrival rate of a continuous lognormal process having the same mean and standard deviation. The tests are conducted over burst distribution of our proposed traffic model discussed in section V. Figure 5 motivates the use of MPTCP in data center networks for improving throughput. Figure 3 is tested using the lognormal distribution with mean μ=1.8 and standard deviation σ=1, corresponding to an arrival rate of 7.12/tu in BCube topology. Figure 3 shows the throughput comparison between TCP (K = 1) and MPTCP (K = 2, 3, 4), where K is the number of paths (i.e., number of subflows) used by each MPTCP connection. It can be observed that, MPTCP gives much higher throughput as compared to single path TCP. It can also be observed that MPTCP performs better with increasing number of paths. Similar results were achieved for FatTree topology.

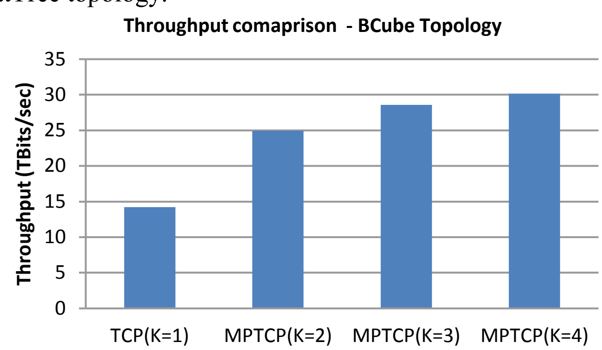

Figure 5. Arrival Rate /μs = 7.12, W=64

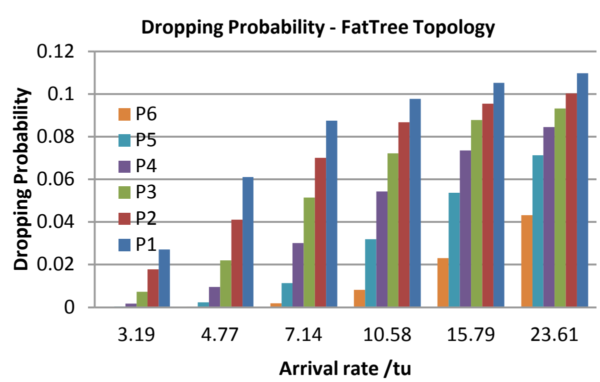

Figure 6. Variable arrival rate, W=64

Figure 6 shows the ability of QAMO-SDN algorithm to achieve QoS differentiation when tested for bursts of various sizes and priority levels as proposed in our traffic model. The dropping probability comparison for six priority levels is shown with increasing load in a FatTree topology. For lognormal traffic, the mean values used in this test are from μ=1 to μ=3 and standard deviation σ=1. It can be observed that the algorithm achieves substantial QoS differentiation for all priority levels. For example, P6 being the highest priority level, experiences the least dropping at all values of input load. Similar results were achieved for BCube topology.

Figure 7 shows the average throughput comparison of TCP, MPTCP (K=4), QAMO and QAMO-SDN. The lognormal mean values used in this test are from μ=0.5 to μ=1.75 and standard deviation σ=1. It can be observed that QAMO and MPTCP (K=4) both performs much better than standard TCP. The throughput of QAMO is slightly less than MPTCP (K=4) at small values of input load while the difference in throughput becomes less at higher loads. QAMO-SDN utlises the available bandwidth better hence there is an improvement in QAMO-SDN compared to QAMO. The reason for QAMO's and QAMO-SDNs' degraded throughput is its preferential treatment for higher priority bursts, which are mostly very small in size.

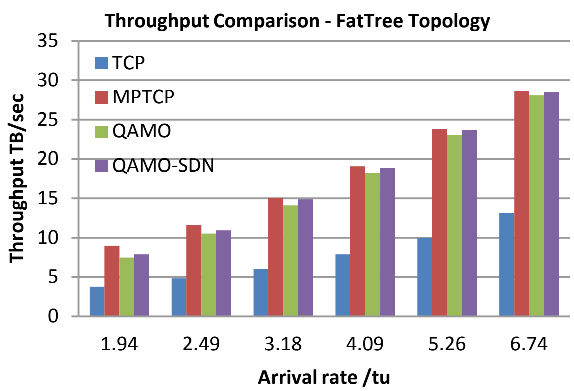

Figure 7. Variable arrival rate, W=64

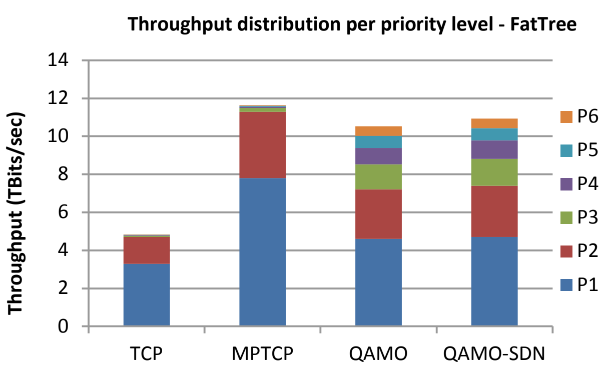

Figure 8. Arrival Rate /μs = 2.49, W=64

Figure 8 provides deeper analysis of throughput breakdown in terms of burst priorities at one of the loads from Figure 5, specifically at arrival rate = 2.49 bursts/tu. The lognormal mean in Figure 6 is μ=0.75 and standard deviation σ=1. It can be observed that in TCP and MPTCP the greatest share of throughput is achieved by low priority background traffic, giving less importance to the time sensitive foreground flows in the absence of QoS provisioning. The throughput of QAMO and QAMO-SDN is well distributed between high priority (foreground) and low priority (background) traffic. Hence, the slight degradation of throughput compared to MPTCP is acceptable for achieving better share of network resources for more critical traffic in data centers. QAMO-SDN achieves better throughput than QAMO due to SDN architecture.

VII. CONCLUSION

In this paper we have shown a possible architecture of the Software defined optical network employing newly emerging transport protocol MPTCP over OBS networks and extended QoS provisioning algorithm for SDN in cloud datacenter. We have seen that MPTCP improves the throughput and reliability in data center networks by parallel transmission on multiple paths. We have presented and evaluated QoS-aware MPTCP over OBS for software defined optical networks (QAMO-SDN) scheme to provide service differentiation in data center traffic. QAMO-SDN provides tangible QoS differentiation to bursts of various classes without impacting the throughput of the system. QAMO-SDN is also an adaptive and self configurable scheme that changes its dynamics based on current network feedback making it applicable in software defined networks (SDN) for future datacenters. QAMO-SDN performs better than our previously proposed QAMO. It must also be noted that the slight improvement in throughput is not the only benefit of SDN architecture over standard QAMO. The motivation to use software defined architecture lies in its simplicity, predictability, ease of network management through a central control and scaling and will continue to grow in the future.

References

[1] Y. Chen, S. Jain, V. K. Adhikari, Z.-L. Zhang, and K. Xu, "A first look at inter-data center traffic characteristics via yahoo! datasets," in *INFOCOM, 2011 Proceedings IEEE*, 2011, pp. 1620-1628.

[2] A. Ghosh, S. Ha, E. Crabbe, M. Chiang, and J. Rexford, "Scalable Multi-Class Traffic Management in Data Center Backbone Networks," *under submission*.

[3] S. T. a. M. Bassiouni, "Performance Evaluation of MPTCP over Optical Burst Switching in Data Centers," presented at the ITS Brazil 2014, São Paulo, Brazil, 2014.

[4] M. Channegowda, R. Nejabati, and D. Simeonidou, "Software-defined optical networks technology and infrastructure: enabling software-defined optical network operations [Invited]," *Optical Communications and Networking, IEEE/OSA Journal of,* vol. 5, pp. A274-A282, 2013.

[5] D. Li, Y. Shang, and C. Chen, "Software Defined Green Data Center Network with Exclusive Routing."

[6] N. McKeown, "Software-defined networking," *INFOCOM keynote talk,* 2009.

[7] C. Monsanto, J. Reich, N. Foster, J. Rexford, and D. Walker, "Composing Software Defined Networks," in *NSDI*, 2013, pp. 1-13.

[8] S. Sezer, S. Scott-Hayward, P.-K. Chouhan, B. Fraser, D. Lake, J. Finnegan*, et al.*, "Are we ready for SDN? Implementation challenges for software-defined networks," *Communications Magazine, IEEE,* vol. 51, 2013.

[9] A. Tootoonchian, S. Gorbunov, Y. Ganjali, M. Casado, and R. Sherwood, "On controller performance in software-defined networks," in *USENIX Workshop on Hot Topics in Management of Internet, Cloud, and Enterprise Networks and Services (Hot-ICE)*, 2012.

[10] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 2010, pp. 267-280.

[11] S. Ranjan, J. Rolia, H. Fu, and E. Knightly, "Qos-driven server migration for internet data centers," in *Quality of Service, 2002. Tenth IEEE International Workshop on*, 2002, pp. 3-12.

[12] Y. Song, H. Wang, Y. Li, B. Feng, and Y. Sun, "Multi-tiered on-demand resource scheduling for VM-based data center," in *Proceedings of the 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid*, 2009, pp. 148-155.

[13] C.-Y. Hong, M. Caesar, and P. Godfrey, "Finishing flows quickly with preemptive scheduling," *ACM SIGCOMM Computer Communication Review,* vol. 42, pp. 127-138, 2012.

[14] S. Liu, H. Xu, and Z. Cai, "Low Latency Datacenter Networking: A Short Survey," *arXiv preprint arXiv:1312.3455,* 2013.

[15] D. Zats, T. Das, P. Mohan, D. Borthakur, and R. Katz, "DeTail: reducing the flow completion time tail in datacenter networks," *ACM SIGCOMM Computer Communication Review,* vol. 42, pp. 139-150, 2012.

[16] C. Wilson, H. Ballani, T. Karagiannis, and A. Rowtron, "Better never than late: Meeting deadlines in datacenter networks," in *ACM SIGCOMM Computer Communication Review*, 2011, pp. 50-61.

[17] Y. Chen, M. Hamdi, and D. H. Tsang, "Proportional QoS over OBS networks," in *Global Telecommunications Conference, 2001. GLOBECOM'01. IEEE*, 2001, pp. 1510-1514.

[18] Q. Zhang, V. M. Vokkarane, B. Chen, and J. P. Jue, "Early drop scheme for providing absolute QoS differentiation in optical burst-switched networks," in *High Performance Switching and Routing, 2003, HPSR. Workshop on*, 2003, pp. 153-157.

[19] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *Selected Areas in Communications, IEEE Journal on,* vol. 18, pp. 2062-2071, 2000.

[20] L. Peng, C.-H. Youn, W. Tang, and C. Qiao, "A novel approach to optical switching for intradatacenter networking," *Lightwave Technology, Journal of,* vol. 30, pp. 252-266, 2012.

[21] A. Lazzez, N. Boudriga, and M. S. Obaidat, "Improving TCP QoS over OBS networks: A scheme based on optical segment retransmission," in *Performance Evaluation of Computer and Telecommunication Systems, 2008. SPECTS 2008. International Symposium on*, 2008, pp. 233-240.

[22] B. Shihada, P.-H. Ho, and Q. Zhang, "A novel congestion detection scheme in TCP over OBS networks," *Journal of Lightwave Technology,* vol. 27, pp. 386-395, 2009.

[23] Q. Zhang, V. M. Vokkarane, Y. Wang, and J. P. Jue, "Analysis of TCP over optical burst-switched networks with burst retransmission," in *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*, 2005, pp. 6 pp.-1983.

[24] C. P. C. Raiciu Sebastien Barre, Adam Greenhalgh, Damon Wischik, and Mark Handley, "Improving Data center Performance and Robustness with Multipath TCP," presented at the ACM SIGCOMM 2011, Toronto, Canada, 2011.

[25] C. Raiciu, C. Pluntke, S. Barre, A. Greenhalgh, D. Wischik, and M. Handley, "Data center networking with multipath TCP," in *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, 2010, p. 10.

[26] S. Gringeri, N. Bitar, and T. J. Xia, "Extending software defined network principles to include optical transport," *Communications Magazine, IEEE,* vol. 51, pp. 32-40, 2013.

[27] N. Cvijetic, "OFDM for next-generation optical access networks," *Lightwave Technology, Journal of,* vol. 30, pp. 384-398, 2012.

[28] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching networks," *Journal of lightwave technology,* vol. 18, p. 2019, 2000.

[29] M. Y. Sowailem, D. V. Plant, and O. Liboiron-Ladouceur, "Implementation of optical burst switching in data centers," in *Photonics Conference (PHO), 2011 IEEE*, 2011, pp. 445-446.

[30] E. Exposito, M. Gineste, L. Dairaine, and C. Chassot, "Building self-optimized communication systems based on applicative cross-layer information," *Computer Standards & Interfaces,* vol. 31, pp. 354-361, 2009.

[31] C. Diop, G. Dugué, C. Chassot, E. Exposito, and J. Gomez, "QoS-aware and autonomic-oriented multi-path TCP extensions for mobile and multimedia applications," *International Journal of Pervasive Computing and Communications,* vol. 8, pp. 306-328, 2012.

[32] C. Diop, G. Dugué, C. Chassot, and E. Exposito, "QoS-aware multipath-TCP extensions for mobile and multimedia applications," in *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia*, 2011, pp. 139-146.

[33] S. Tariq and M. A. Bassiouni, "Hop-count fairness-aware protocols for improved bandwidth utilization in WDM burst-switched networks," *Photonic Network Communications,* vol. 25, pp. 35-46, 2013.

[34] S. Tariq, M. Bassiouni, and G. Li, "Improving Fairness of OBS Routing Protocols in Multimode Fiber Networks," presented at the Computing, Networking and Communications (ICNC), 2013, 2013.

[35] T. Benson, A. Anand, A. Akella, and M. Zhang, "Understanding data center traffic characteristics," *ACM SIGCOMM Computer Communication Review,* vol. 40, pp. 92-99, 2010.

[36] M. Pustisek, I. Humar, and J. Bester, "Empirical analysis and modeling of peer-to-peer traffic flows," in *Electrotechnical Conference, 2008. MELECON 2008. The 14th IEEE Mediterranean*, 2008, pp. 169-175.

[37] X. Gao and M. A. Bassiouni, "Improving fairness with novel adaptive routing in optical burst-switched networks," *Journal of Lightwave Technology,* vol. 27, pp. 4480-4492, 2009.

[38] S. Saha, J. S. Deogun, and L. Xu, "Hyscaleii: A high performance hybrid optical network architecture for data centers," in *Sarnoff Symposium (SARNOFF), 2012 35th IEEE*, 2012, pp. 1-5.

[39] M. Alizadeh, A. Greenberg, D. A. Maltz, J. Padhye, P. Patel, B. Prabhakar*, et al.*, "Data center tcp (dctcp)," *ACM SIGCOMM Computer Communication Review,* vol. 40, pp. 63-74, 2010.