QAMO: QoS Aware Multipath-TCP Over Optical Burst Switching in Data Centers

Sana Tariq and Mostafa Bassiouni
Department of Elec. Eng. & Computer Science, University of Central Florida, Orlando, U.S.A.

Keywords: MPTCP, All-optical Network, Data Center, TCP, OBS, QoS.

Abstract: The rapid advancement in cloud computing is leading to a promising future for shared data centers hosting diverse applications. These 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 to provide potential service differentiation to client's individual flows. This paper addresses two important issues in shared data centers: bandwidth efficiency and service differentiation based on QoS (Quality of Service). We first evaluate the multipath-TCP (MPTCP) protocol over an OBS (Optical burst switching) network for improved bandwidth utilization in dense interconnect datacenter networks. We next present a simple and efficient 'QoS aware MPTCP over OBS' (QAMO) algorithm in datacenters. Our experimental results show that Multipath-TCP improves throughput over an OBS network while the QAMO algorithm achieves tangible service differentiation without impacting the throughput of the system.

1 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. Over the past decade, considerable attention has been given to different areas of cloud computing e.g., efficient sharing of computational resources, virtualization, scalability and security. However, less attention has been paid to network management and QoS (Quality-of-Service) provisioning in datacenters. The inability of today's cloud technologies to provide dependable and predictable services is a major showstopper for the widespread adoption of the cloud paradigm (Rygielski and Kounev, 2013).

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 (Chen et al., 2011). 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 (Ghosh et al., 2013).

In this paper, we first evaluate the performance of MPTCP over OBS for datacenter networks and compare the performance of TCP with MPTCP under different network loads and topologies. We next present and evaluate a QoS provisioning algorithm called QAMO, 'QoS aware MPTCP over OBS'. To our knowledge, this is the first research report that provides QoS provisioning algorithm for service differentiation using MPTCP over OBS in datacenters.

The rest of the paper is organized as follows. In section 2, we review previous work. In section 3, we describe our networking model that uses MPTCP protocol over an optical burst switching network for data centers. In section 4, we present 'QoS-aware MPTCP over OBS' (QAMO) scheme. Simulation details are discussed in section 5 and the performance analysis and simulation results are given in Section 6. We conclude the paper in Section 7.

Tariq S. and Bassiouni M..
QAMO: QoS Aware Multipath-TCP Over Optical Burst Switching in Data Centers.
DOI: 10.5220/0005047800560063
In *Proceedings of the 5th International Conference on Optical Communication Systems* (OPTICS-2014), pages 56-63
ISBN: 978-989-758-044-4

2 PREVIOUS WORK

Future data center consumers will require quality of service QoS as a fundamental feature. There have been some research studies on traffic modeling, network resource management and QoS provisioning in data centers (Chen et al., 2011), (Benson et al., 2010a) and (Ranjan et al., 2002). Ranjan, et. al., studied the problem of QoS guarantees in data-center environments in (Ranjan et al., 2002). However, this work is not suitable for highly loaded shared data-centers with computationally intensive applications due to the two sided nature of communication. The work in (Song et al., 2009) 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. A flow scheduling protocol called Preemptive Distributed Quick (PDQ) in (Hong et al., 2012) was, designed to complete flows quickly by emulating a shortest job first algorithm and giving priority to the short flows. Similarly research in (Liu et al., 2013) proposed taxonomy to categorize existing works based on three main techniques: reducing queue length, prioritizing mice flows, and exploiting multi-path. DeTail (Zats et al., 2012) is an idea of cross-layer network stack aiming to improve the tail of completion time for delay-sensitive flows. A deadline-aware control protocol was presented in (Wilson et al., 2011), 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 datacenter networks. The techniques of giving priority to mice flows and exploiting multiple of newer datacenter topologies have proved to be effective and efficient means of achieving service differentiation in datacenters.

However none of the presented QoS provisioning schemes explored the combined approach of using multiple paths of newly emerging transport protocols such as Multi-path TCP (MPTCP) with wavelength reservation at network layer in optical domain to achieve service differentiation. Our proposed scheme combines the flexibility of bandwidth reservation at two levels to achieve QoS in datacenters that marks its novelty in approach. There is rich research on QoS schemes in optical burst switching for wide area networks (Chen et al., 2001), (Zhang et al., 2003), (Yoo et al., 2000) and (Akar et al., 2010). 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 (Peng et al., 2012). TCP is the most dominant transport layer protocol in internet and TCP over OBS has been extensively studied (Lazzez et al., 2008), (Shihada et al., 2009) and (Zhang et al., 2005). Multipath-TCP (MPTCP) has been shown to provide significant improvement in throughput and reliability in electronic packet switched networks in data centers (Raiciu et al., 2011) and (Raiciu et al., 2010). However, MPTCP has not been studied in the context of OBS networks before. In this paper we will develop a QoS provisioning scheme for data center networks using MPTCP over OBS and evaluate its performance.

3 NETWORK MODEL

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). The control information is electronically processed at each optical router while the payload is transmitted all-optically with full transparency through the lightpath. The wavelength reservation protocol plays a crucial role in the burst transmission and we have used just-in-time (JIT) (Wei and McFarland, 2000) for its simplicity. The necessary hardware level modifications of optical switches for supporting OBS in data centers have been discussed in (Sowailem et al., 2011), and will not be repeated in this paper.

4 QoS AWARE MPTCP OVER OBS ALGORITHM

Our proposed algorithm QoS aware MPTCP over OBS called QAMO 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. QAMO algorithm does that as follows:

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 $P_1, P_2, P_3, \ldots, P_{max}$ while P_{max} is the highest priority level in the bursts. We next define the number of allocated paths k_{curr} for the burst of a particular priority level as follows in equation 1.

$$k_{curr} = \lceil K \cdot P \rceil \quad (1)$$

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 $P_{curr} = 0.5 \cdot P_{max}$, 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 equation 2

$$Wavelengthsearchsize = \lceil W \cdot P \rceil \quad (2)$$

At wavelength reservation stage in OBS, equation 2 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.

In algorithm 1, 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. 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. QAMO algorithm should only be active when the network is congested and the received traffic exhibits difference

Algorithm 1 : QAMO (QoS Aware MPTCP over OBS) Algorithm.

```
Input:
P = P_curr / P_max
K = maximum number of paths
W = maximum number of wavelengths
w_curr = current wavelength reserved for current burst
N_k = vector of all nodes on path k
k_curr = paths allocated to the current burst
burst_curr = current burst

Algorithm:
for each k in ⌈K · P⌉ do
    generate k lightpaths by making k concurrent function calls to generateLightPath()
end for

generateLightpath(path k)
Initialize w_curr
for each n in N_k do
    if n is the destination node then
        break
    end if
    if n is the source node then
        for each w in ⌈W · P⌉ do
            if w is free then
                reserve w for burst_curr at n
                w_curr = w break
            end if
        end for
    else
        if w_curr is free at n then
            reserve w_curr for burst_curr at n continue
        end if
        for each w in ⌈W · P⌉ do
            if w is free then
                reserve w for burst_curr at n
                w_curr = w break
            end if
        end for
    end if
    if no free wavelength at n then
        return (error) Search failed at node n
    end if
end for
return (success)
```

between relative priority levels.

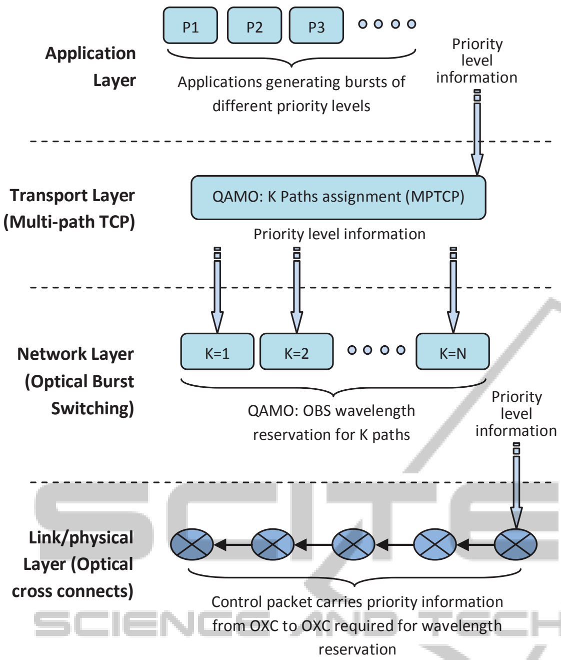

Figure 1: QAMO's cross-layer design: Changes to the protocol stack and the burst priority level information flow.

As shown in figure 1, we assume that QAMO algorithm has access to available information about QoS requirements of different bursts to process them correctly. AT MPTCP layer this capability may be implemented using a specific interface such as the Implicit Packet Meta Header (IPMH) promoted in (Exposito et al., 2009). Because of MPTCP IPMH interface (Diop et al., 2012) and (Diop et al., 2011), it is possible to assign priority levels for different flows and gather priority information for each type of flow at a particular end host. Under QAMO scheme, when the burst priority level information is received at MPTCP layer, QAMO initiates k lightpath requests to the OBS network layer. 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 easily be passed on to lower layer and then from one one OXC to the next 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, to find a free wavelength requires minor modification to the standard JIT channel allocation scheme. The adjustable search in a smaller space of $\lceil W \cdot P \rceil$ for wavelengths actually leads to a smaller average search time.

The QAMO 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.

5 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 (Tariq and Bassiouni, 2013) and (Tariq et al., 2013). 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 (Benson et al., 2010a) and (Benson et al., 2010b). The lognormal distribution is also considered to be the most fitted distribution for modeling various categories of internet traffic including TCP (Pustisek et al., 2008). 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 (Tariq and Bassiouni, 2013) and (Gao and Bassiouni, 2009).

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 (Saha et al., 2012). Each node can have a certain maximum number W of allowed wavelengths. Arrival rate/μs 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 (Benson et al., 2010a). The medium sized and longer flows constitute background traffic and may belong to different priority levels (Alizadeh et al., 2011). To represent these scenarios of data center mixed traffic, we have used variable burst sizes in different ranges with uniform distribution within each range (Alizadeh et al., 2011).

Short burst sizes: $S_{min} = 5$ Kbits to $S_{max} = 20$ Kbits
Medium burst sizes: $S_{min} = 200$ Kbits to $S_{max} = 1$ Mbits
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 (Chen et al., 2011), (Benson et al., 2010a), (Benson et al., 2010b) and (Alizadeh et al., 2011). 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].

6 RESULTS AND DISCUSSION

The topologies used in our simulation tests are FatTree with 36 nodes and BCube with 24 nodes as shown in Figure 2 and Figure 3. 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 2 in the bottom layer.

The BCube network shown in Figure 3 has 16 relaying hosts in the middle layer. The network diameter in the 24-node BCube network is 4 hops.

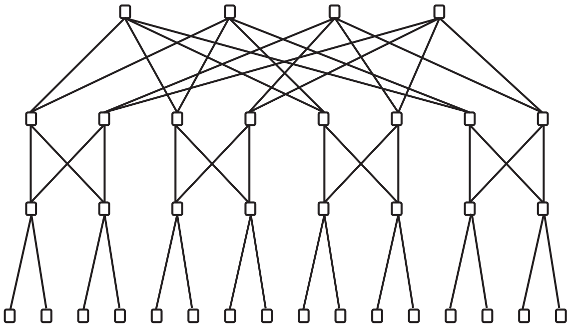

Figure 2: FatTree topology used for simulation.

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 5.

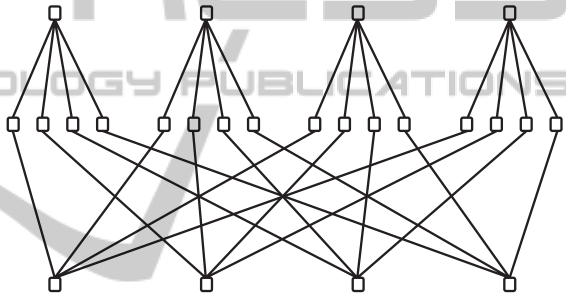

Figure 3: BCube topology used for simulation.

Figure 4 motivates the use of MPTCP in data center networks for improving throughput. Figure 4 is tested using the lognormal distribution with mean $\mu = 1.8$ and standard deviation $\sigma = 1$, corresponding to an arrival rate of $7.12/tu$ in BCube topology. Figure 4 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 shows the ability of QAMO 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 $\mu = 1$ to $\mu = 3$ and standard deviation $\sigma = 1$. It can be observed that the algorithm achieves substantial QoS differentiation for all priority levels. For example, $P6$ being the highest

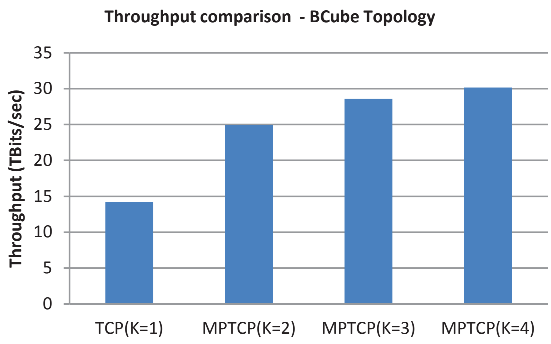

Figure 4: Arrival Rate $/\mu s = 7.12$, $W = 64$.

priority level, experiences the least dropping at all values of input load. Similar results were achieved for BCube topology.

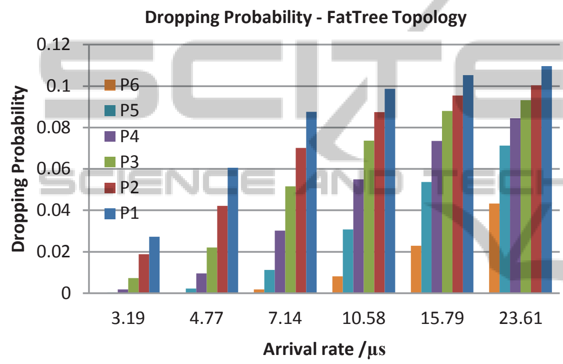

Figure 5: Variable arrival rate, $W = 64$.

Figure 6 shows the average throughput comparison of TCP, MPTCP $(K = 4)$ and QAMO. The lognormal mean values used in this test are from $\mu = 0.5$ to $\mu = 1.75$ and standard deviation $\sigma = 1$. 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. The reason for QAMO's degraded throughput is its preferential treatment for higher priority bursts, which are mostly very small in size.

Figure 7 provides deeper analysis of throughput breakdown in terms of burst priorities at one of the loads from Figure 6, specifically at arrival rate = 2.49 $bursts/tu$. The lognormal mean in Figure 7 is $\mu = 0.75$ and standard deviation $\sigma = 1$. 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 is well distributed between high priority (foreground) and low priority (background) traffic. Hence, the slight degradation of QAMO throughput compared

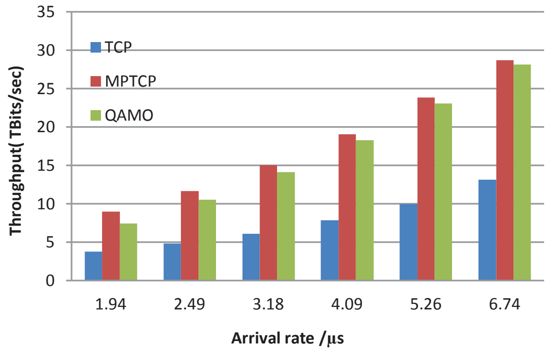

Figure 6: Variable arrival rate, $W = 64$.

to MPTCP is acceptable for achieving better share of network resources for more critical traffic in data centers.

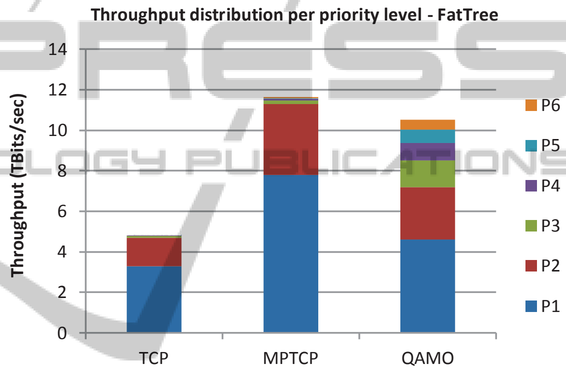

Figure 7: Arrival Rate $/\mu s = 2.49$, $W = 64$.

7 CONCLUSIONS

In this paper we have shown the benefits of the newly emerging transport protocol MPTCP over OBS networks in data centers. 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 (QAMO) scheme to provide service differentiation in data center traffic. QAMO algorithm provides tangible QoS differentiation to bursts of various classes without impacting the throughput of the system. For future work, we plan to improve basic QAMO scheme to an adaptive and self configurable algorithm that can change its dynamics based on current network feedback. These extensions will make it applicable in software defined networks (SDN) for future datacenters.

REFERENCES

Akar, N., Karasan, E., Vlachos, K. G., Varvarigos, E. A., Careglio, D., Klinkowski, M., and Solé-Pareta, J. (2010). A survey of quality of service differentiation mechanisms for optical burst switching networks. *Optical Switching and Networking*, 7(1):1–11.

Alizadeh, M., Greenberg, A., Maltz, D. A., Padhye, J., Patel, P., Prabhakar, B., Sengupta, S., and Sridharan, M. (2011). Data center tcp (dctcp). *ACM SIGCOMM computer communication review*, 41(4):63–74.

Benson, T., Akella, A., and Maltz, D. A. (2010a). Network traffic characteristics of data centers in the wild. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 267–280. ACM.

Benson, T., Anand, A., Akella, A., and Zhang, M. (2010b). Understanding data center traffic characteristics. *ACM SIGCOMM Computer Communication Review*, 40(1):92–99.

Chen, Y., Hamdi, M., and Tsang, D. H. (2001). Proportional qos over obs networks. In *Global Telecommunications Conference, 2001. GLOBECOM'01. IEEE*, volume 3, pages 1510–1514. IEEE.

Chen, Y., Jain, S., Adhikari, V. K., Zhang, Z.-L., and Xu, K. (2011). A first look at inter-data center traffic characteristics via yahoo! datasets. In *INFOCOM, 2011 Proceedings IEEE*, pages 1620–1628. IEEE.

Diop, C., Dugué, G., Chassot, C., and Exposito, E. (2011). Qos-aware multipath-tcp extensions for mobile and multimedia applications. In *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia*, pages 139–146. ACM.

Diop, C., Dugué, G., Chassot, C., Exposito, E., and Gomez, J. (2012). Qos-aware and autonomic-oriented multi-path tcp extensions for mobile and multimedia applications. *International Journal of Pervasive Computing and Communications*, 8(4):306–328.

Exposito, E., Gineste, M., Dairaine, L., and Chassot, C. (2009). Building self-optimized communication systems based on applicative cross-layer information. *Computer Standards & Interfaces*, 31(2):354–361.

Gao, X. and Bassiouni, M. A. (2009). Improving fairness with novel adaptive routing in optical burst-switched networks. *Journal of Lightwave Technology*, 27(20):4480–4492.

Ghosh, A., Ha, S., Crabbe, E., and Rexford, J. (2013). Scalable multi-class traffic management in data center backbone networks. *Selected Areas in Communications, IEEE Journal on*, 31(12):2673–2684.

Hong, C.-Y., Caesar, M., and Godfrey, P. (2012). Finishing flows quickly with preemptive scheduling. *ACM SIGCOMM Computer Communication Review*, 42(4):127–138.

Lazzez, A., Boudriga, N., and Obaidat, M. S. (2008). 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*, pages 233–240. IEEE.

Liu, S., Xu, H., and Cai, Z. (2013). Low latency datacenter networking: A short survey. *arXiv preprint arXiv:1312.3455*.

Peng, L., Youn, C.-H., Tang, W., and Qiao, C. (2012). A novel approach to optical switching for intradatacenter networking. *Lightwave Technology, Journal of*, 30(2):252–266.

Pustisek, M., Humar, I., and Bester, J. (2008). Empirical analysis and modeling of peer-to-peer traffic flows. In *Electrotechnical Conference, 2008. MELECON 2008. The 14th IEEE Mediterranean*, pages 169–175. IEEE.

Raiciu, C., Barre, S., Pluntke, C., Greenhalgh, A., Wischik, D., and Handley, M. (2011). Improving datacenter performance and robustness with multipath tcp. In *ACM SIGCOMM Computer Communication Review*, volume 41, pages 266–277. ACM.

Raiciu, C., Pluntke, C., Barre, S., Greenhalgh, A., Wischik, D., and Handley, M. (2010). Data center networking with multipath tcp. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, page 10. ACM.

Ranjan, S., Rolia, J., Fu, H., and Knightly, E. (2002). Qos-driven server migration for internet data centers. In *Quality of Service, 2002. Tenth IEEE International Workshop on*, pages 3–12. IEEE.

Rygielski, P. and Kounev, S. (2013). Network virtualization for qos-aware resource management in cloud data centers: A survey. *Praxis der Informationsverarbeitung und Kommunikation*, 36(1):55–64.

Saha, S., Deogun, J. S., and Xu, L. (2012). Hyscaleii: A high performance hybrid optical network architecture for data centers. In *Sarnoff Symposium (SARNOFF), 2012 35th IEEE*, pages 1–5. IEEE.

Shihada, B., Ho, P.-H., and Zhang, Q. (2009). A novel congestion detection scheme in tcp over obs networks. *Journal of Lightwave Technology*, 27(4):386–395.

Song, Y., Wang, H., Li, Y., Feng, B., and Sun, Y. (2009). 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*, pages 148–155. IEEE Computer Society.

Sowailem, M. Y., Plant, D. V., and Liboiron-Ladouceur, O. (2011). Implementation of optical burst switching in data centers. *IEEE PHO, WG4*.

Tariq, S., Bassiouni, M., and Li, G. (2013). Improving fairness of obs routing protocols in multimode fiber networks. In *Proceedings of the 2013 International Conference on Computing, Networking and Communications (ICNC)*, pages 1146–1150. IEEE Computer Society.

Tariq, S. and Bassiouni, M. A. (2013). Hop-count fairness-aware protocols for improved bandwidth utilization in wdm burst-switched networks. *Photonic Network Communications*, 25(1):35–46.

Wei, J. Y. and McFarland, R. I. (2000). Just-in-time signaling for wdm optical burst switching networks. *Journal of lightwave technology*, 18(12):2019.

Wilson, C., Ballani, H., Karagiannis, T., and Rowtron, A. (2011). Better never than late: Meeting deadlines in datacenter networks. In *ACM SIGCOMM Computer Communication Review*, volume 41, pages 50–61. ACM.

Yoo, M., Qiao, C., and Dixit, S. (2000). Qos performance of optical burst switching in ip-over-wdm networks. *Selected Areas in Communications, IEEE Journal on*, 18(10):2062–2071.

Zats, D., Das, T., Mohan, P., Borthakur, D., and Katz, R. (2012). Detail: reducing the flow completion time tail in datacenter networks. *ACM SIGCOMM Computer Communication Review*, 42(4):139–150.

Zhang, Q., Vokkarane, V. M., Chen, B., and Jue, J. P. (2003). Early drop scheme for providing absolute qos differentiation in optical burst-switched networks. In *High Performance Switching and Routing, 2003, HPSR. Workshop on*, pages 153–157. IEEE.

Zhang, Q., Vokkarane, V. M., Wang, Y., and Jue, J. P. (2005). Analysis of tcp over optical burst-switched networks with burst retransmission. In *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*, volume 4, pages 6–pp. IEEE.