Optimization of Bandwidth Utilization in Data Center Network with SDN

Yaofang Li[1,2,a], Bin Wu[3,b] Jie Xiao[4,c], Chunxia Dai[5,d]
[1]Computing Center, Tianjin Chengjian University, Tianjin, 300384, P. R. China.
[2]School of Computer Science and Technology, Tianjin University, Tianjin, 300072, P. R. China
[3]School of Computer Science and Technology, Tianjin University, Tianjin, 300072, P. R. China
[4]School of Computer Science and Technology, Tianjin University, Tianjin, 300072, P. R. China
[5]Computing Center, Tianjin Chengjian University, Tianjin, 300384, P. R. China
[a]emails: yaofangli@163.com, [b]emails:, binwu.tju@gmail.com, [c]emails: jiexiao001@gmail.com, [d]emails: daichx@tcu.edu.cn

Keywords: Data center network(DCN), delay, capacity utilization, SDN (Software-defined Networking).

Abstract. Data-center network (DCN) plays a fundamental role in cloud computing. As applications grow fastly, capacity utilization of data-center network becomes a big challenge to cloud service, especially when user's requests are unbalanced or data-center works at a peak time when there are amount of demands need to be handled. Software-defined Networking (SDN) is an efficient technology to manage network utilization. In this paper, we apply SDN controller for service routing to increase bandwidth utilization of DCN and meanwhile reduce delay of end-users, by proposing a heuristic algorithm for SDN. Numerical experiments demonstrate the good performance of the proposed algorithm.

Introduction

A data center network (DCN) [1-2] generally consists of a very large number of servers hosted in a factory-scale warehouse. These severs provide effective services to hundreds of millions of end-users. The enormous and growing demand of these applications has motivated service providers to deploy geographically distributed datacenters for both reliability and performance reasons. The resulting deployment of application leads to a particular important request allocation problem, which means that massive end-users' requests across the wide area must be directed to an appropriate datacenter [3].

Normally, such request allocation is able to be completed with a simple method that allocates each request to the closest datacenter [4]. This will lead to another urgent problem that the closet datacenter may be overloaded to be outworked in a extreme case. So with the rapid increase of traffic in DCN and the growing demands from users, how to efficiently manage network resources became an urgent need that attracted numerous researchers' attention.

Many people have been devoted in improving QoS of DCN services by achieving high utilization of DCN internal bandwidth. Different DCN topologies have been proposed to improve scalability and efficiency of communication bandwidth [5-10]. Some studies try to minimize network operational costs or average latency by caching files at different locations of the network [11-13]. [14-15] proposed a scalable architecture by properly placing server racks to minimize system cost. Nevertheless, the above solutions cannot solve the problem that how to manage service routing in real-time by taking current DCN status and user demands into account for achieving a balanced bandwidth utilization.

SDN (software defined network) [16] can perfectly cope with this situation. It allows design and implementation of more flexible and programmable networks [17-19]. SDN is featured by separating control and data planes with a centralized control plane to track global information of network status [20]. It can be applied to various networks such as DCN. The work [21] gives the first try to combine SDN with DCN in 2009, and since then more and more novel ideas have been proposed on this.

Specifically, [22] focuses on resource management in DCNs with SDN control. It indicates that SDN can bring network awareness to DCN applications, and meanwhile make DCN traffic routing more efficiently. Similiar to [22], [23] proposes a dynamic load balancing method for DCN based on SDN. [24-27] try to increase DCN bandwidth utilization by SDN architecture design. Authors in [28] presents a Nash bargaining method to achieve high bandwidth utilization (from service provider's perspective) and low latency (from end-user's perspective) by using SDN controller, but bandwidth utilizations of different providers are simply multiplied together. Though there are various existing SDN based solutions for improving DCN network utilization, they have not taken routing into account under certain failure probabilities of DCN internal links. While in our early research we haven't take capacity utilization into account that will be result in an imbalance using in all data centers.

In this paper, we combine SDN with DCN to balance capacity utilization of all data centers and meanwhile decrease service delay of end-users, as well as improving bandwidth utilization of DCN. Based on our earlier work [29], a routing algorithm is designed for SDN controller to program network resources, such that bandwidth utilization of different service providers can be balanced in an even state. We first formulate an optimization model, and then design a heuristic routing algorithm by taking transmission delay and link failure probability into account.

The rest of the paper is organized as follows. We describe system model and problem in Section II. Section III formulates the optimization model and proposes the routing algorithm. Section IV presents numerical results. We conclude the paper in Section V.

System Model and Problem Description

A. System Model

In a scale DCN, there will be thousands or millions demands asking to be served. These demands will be satisfied by the nearest data center in a normal way in order to decrease the transmission cost. When the resources of DCN are sufficient or the demands are not adequate it does save transmission cost and total cost of DCN. But with the number of web users increasing, the resources of data center cannot develop synchronously with them. This will lead to a phenomenon that some data centers cannot satisfy their nearest users while other data centers may be idle at some peak time. In this situation, the users face two choices which is either keep waiting for its nearest datacenter until it is not busy or turning to ask for services from other datacenter. Which choice is ok depending on whose total cost(transmission cost and delay cost) is small. We examine this situation and try to solve the balancing of demands and resources. Our aim is to optimize the bandwidth utilization of datacenters in DCN with the help of SDN. The system model, named of OBU-SDN(Optimizing Bandwidth Utilization by SDN) is illustrated in Fig.1, where the SDN controller connects users and DCN. The DCN consists of several DCs, each with a different bandwidth capacity. Users deliver their demands to SDN controller rather than directly to data centers, and then SDN controller orders the DCN to feed back the current information of each DCs (bandwidth utilization, jam situation, and transmission delay, etc.). Next, SDN assigns resource to users and decides serving routing based on the information. SDN controller plays as an administrator in the backbone. It makes overall arrangements in the network to minimize resource waste, and meanwhile prevent a particular DC from traffic jam. This also benefits end-users by balancing the workload of different DCs, though end-users may experience a little bit more delay by getting service from a remote DC. In fact, delay can indeed be decreased if the DCs can better serve end-users in a load-balancing manner.

Similar to our previous work [29], we assume that each link entails a certain transmission delay and is with a particular failure probability. The specific values of transmission delay and failure probability of each link can be obtained from engineering practice. Since the values of the above two parameters are both proportional to the corresponding link length, we can integrate them into one cost by scaling factor. We assume that each user has a service request that matches a certain amount of cloud resources, and it is mapped to a particular node in the backbone network. The total

demands of a node equal the sum of all end-user demands at this node. The total capacity of all DCs should be abundant to satisfy the demands of all users..

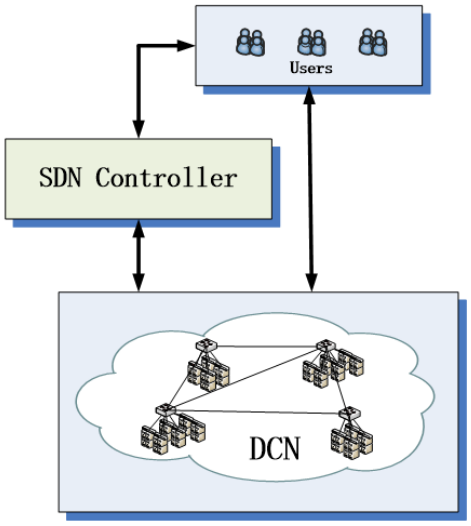

Fig. 1 the DCN architecture, where SDN controller controls the whole network

B. Problem Description

Our objective is to minimize the variance of bandwidth utilization of data centers and the total cost of serving all users with consideration of reducing service failure potential and service delay.

Demands jam may occur at a data center if users are always served by the closest nearby DC, especially at traffic peak time. The unbalanced capacity utilization among different data centers may lead to some DCs become quite busy while others are completely idle, and the former situation will lead to unnecessary user delay and thus poor QoS of cloud service. To increase the overall bandwidth utilization, we adopt SDN controller to balance the bandwidth utilization among all data centers. At peak time, when the nearest datacenter does not have enough capacity, user demands will be transferred to the second nearest data-center. We assume this will happen if the bandwidth utilization of the nearest data-center exceeds the critical threshold value ε. The latter demands can take advantage by this while these traveling demands decrease the waiting queue length and traveling requests' transmission delay may be getting larger thus increasing the total delay. However, this may not always be the case. It is possible that the total cost will decrease if the data center spends more time to handle one demand. The total cost's reduction is owe to that the waiting cost for the nearest data center is much bigger than the transmission cost for the further data center.

Based on the framework described in Section II.A, we need to determine the access route of demands originated from users at each node and allocate corresponding DC resources. Once a unique route is set up between a node and a DC, it is responsible for data exchanging between them. To increase resource utilization and network reliability, we propose an algorithm for SDN to arrange the overall resources to each user (detailed in Section III).

Factors considered for service routing include service cost and delay as well as failure potential of links for reliable service transmissions. Our problem includes the following three aspects: 1) balance bandwidth utilization among all data centers; 2) find proper service routes for external demands; and 3) minimize the overall cost of the system.

Formulation and Algorithm

Bandwidth utilization and transmission cost are taken as our optimization objective. To avoid unbalanced data-center utilization, we minimize the variance of bandwidth utilization among all data centers.

Variance is an effective value to describe if resources are distributed to all members evenly. In DCN, numerous requests search for service from several data centers. The awful situation is unpopular that one data center serves most demands while others are idle. So we do our best to even

the demands to all data centers in order to balance the bandwidth of DCN. We compute the variance of bandwidth of all data centers in DCN to report the utilization level of it.

We try to minimize transmission cost through shortest path algorithm between users and target data center. It is assumed that remote data transmissions across fiber links will lead to transmission cost and delay, and will be subject to link failure probabilities as well. Two cost factors are introduced to integrate delay and failure probability with the system cost. The delay and transmission costs of all traffics are weighted by a cost scaling factor θ. To take service failure potential into account, we define another cost scaling factor η to translate it into a countable cost parameter.

We propose an optimization model and a heuristic algorithm in this section. The former minimizes the total service cost and the variance of bandwidth utilization among all data centers. We introduce two coefficients ρ and φ in our formula to make the value of bandwidth and transmission delay closer to each other. Based on the optimization model, a heuristic algorithm is proposed for SDN to achieve proper service routing control in a given backbone network under a set of predefined service demands.

A. Notation List:

Inputs:

$\boldsymbol{V}$: The set of all nodes in network ***G(V, E)***.

$\boldsymbol{E}$: The set of all directional links in network ***G(V, E)***.

Q: The service capacity of single rack. In this work, it is assumed every rack owns the same capacity.

P_{uv}: Transmission cost of per unit traffic between nodes *u* and *v* in the optical network. In this work, it is assumed as the distance between nodes *u* and *v*, and $P_{uv} = P_{vu}$.

f_{uv}: The failure probability of link (*u*, *v*), $(u, v) \in \boldsymbol{E}$.

r_d: The service demand at a network node, $d \in \boldsymbol{V}$.

η: Cost scaling factor for service failure potential.

θ: Cost scaling factor for taking the inter-node transmission delay of DCN internal traffic into account. $\theta > 1$ and θP_{uv} is the total cost of the DCN internal overhead for per unit internal traffic transmission between nodes *u* and *v*, where the cost of delay counts for $(\theta - 1)P_{uv}$.

$\boldsymbol{V}_s$:The set of nodes placed with DCs. $\boldsymbol{V}_s \in \boldsymbol{V}$.

M_s^*:The number of racks in the DC placed at node s. $s \in \boldsymbol{V}_s$.

α: A predefined constant and $\alpha \geq \sum_{s \in \boldsymbol{V}_s} M_s^*$.

ρ: Scaling factor for total cost.

φ: Scaling factor for the variance of bandwidth utilization.

ε: Critical value of bandwidth utilization of one data center.

Variables:

W_s^d: Nonnegative integer variable. It is the demands at node $d \in \boldsymbol{V}$ that are served by the data center at node $s \in \boldsymbol{V}_s$.

τ: Time for data-center deal with one demand.

Next, we define the total cost (transmission delay, transmission cost, and failure potential) at each link as the *transmission related cost* , and calculate it according to θ and η. We take q_{uv} to denote it and get the following equation:

$$q_{uv} = \theta P_{uv} - \eta \ln(1 - f_{uv}) \quad (1)$$

The first term in function (1) denotes transmission and delay cost by factor θ, and the second term formulates the cost of the service failure potential at each link.

Furthermore, based on the *transmission related cost* of each link, we can calculate the shortest distance between two arbitrary nodes *u* and *v*, and denote it by h_{uv}. The shortest path can also be achieved in accompany with it.

Finally, we define σ as the total bandwidth utilization ratio of all data centers, which is expressed as:

$$\sigma = \frac{\sum_{d \in V} r_d}{Q \sum_{s \in V_s^*} M_s^*} \tag{2}$$

For the sake of cost minimization, the system is destined to choose the shortest path to deliver service. Consequently, different from our earlier research [29], in this work, we combine the variance of bandwidth utilization and total cost as our optimize target.

B. OBU-SDN Formulation

$$minimize \left\{ \mathcal{C} = \rho \sum_{d \in \boldsymbol{V}} \sum_{s \in \boldsymbol{V}_s} W_s^d \, h_{ds} + \varphi \sum_{s \in V_s^*} \left(\frac{\sum_{d \in V} W_s^d}{Q M_s^*} - \sigma \right)^2 \right\} \tag{3}$$

Subject to:

$$\sum_{s \in \boldsymbol{V}_s} W_s^d = r_d \, , \forall d \in \boldsymbol{V} \tag{4}$$

$$\frac{\sum_{d \in V} W_s^d}{Q M_s^*} \leq \varepsilon \leq 1, \ \forall s \in \boldsymbol{V}_s \tag{5}$$

$$\sigma \leq \varepsilon \leq 1 \tag{6}$$

Objective (3) minimizes the link cost and the variance bandwidth utilization in OBU-SDN. Constraint (4) ensures the demand at each node to be severed exactly. Constraint (5) limits the total service amount of a particular data center to its maximum service capacity and must not exceed its critical value. Constraint (6) ensures the total service amount of DCN must be limited to its capacity range.

Equation (3) is our optimize result and how to get the minimize value is another problem. Next, we should decide the serving routing of each user to ensure our optimization formula. We proposed our algorithm as below:

Fig. 2 algorithm of deciding service demand routing

Algorithm: decide serving rout for each node

Numerical Results

In this section, we set up numerous simulations to verify the proposed algorithm based on the network with 30 nodes and 62 links, which is taken from [30]. We assume that there are three data-centers and each has 7,5 and 5 racks respectively. The transmission delay and failure probability of each link are based on our previous work[29].

In order to analyze how the SDN arrange the data-centers bandwidth utilization at the peak time, we examined user requests at 8 time spots to represent different data traffic. And the requests are becoming more and more imbalance. In our simulate DCN, there are 3 data centers, which are distributed at node 5, 13 and 22. In our simulation we assumed the other algorithm named Non-SDN which not combining SDN with DCN architecture compared with our algorithm.

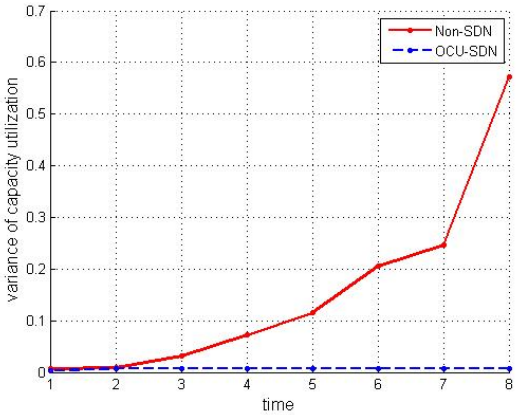

Fig. 3 the change of variance of capacity utilization with 8 kinds of different requests

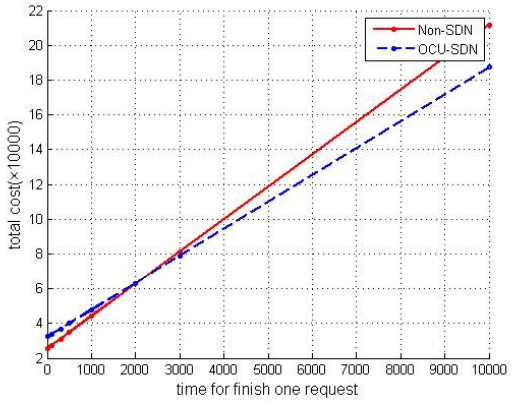

Fig. 4 the change of total cost and variance of capacity utilization with 8 kinds of different requests

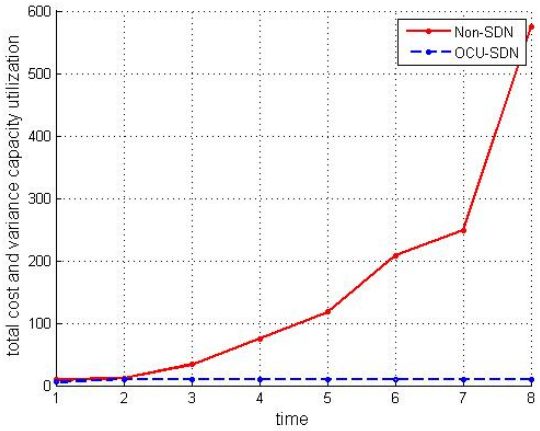

Fig. 5 the change of total cost while serving time for one demand changes

Fig.3 shows the changing in the variance of all data-centers bandwidth utilization under different time. At time 1-2, almost every user can be served by their nearest data center. But at other time, the requests began to concentrate at data center 5 and 22. From Fig.3 we can see the value of Non-SDN is increasing violently because some data centers are overused while others are free seriously. The blue line demonstrates that all data centers work at an even state in our OBU-SDN.

Fig.4 shows how the minimize total result $\mathcal{C}$ of equation (3) changes with different user requests . For better analyze the total value including total cost and the variance, we set the two coefficients φ =1000 and η =0.0001 to help the two items nearer to each other. From this picture, we can obviously conclude that OBU-SDN is preceded than Non-SDN because the value is much smaller than that of Non-SDN.

Fig. 5 indicates how serving time of data center influences the total cost of two DCNs. From Fig.5 we can see that the total cost is increasing as the data center will spend more time to serve one request, and that is quite reasonable. When the processing time is lower, the total cost of Non-SDN is smaller than that of OBU-SDN because Non-SDN users don't need to travel to other data center which further. They just wait and the waiting time can be ignored. But when a data center takes more time to finish one request, users in Non-SDN should spend more time to wait thus result to huge delay cost. Then the advantage of OBU-SDN has evolved slowly.

Fig. 6 lists the serving routes of demands at each node in the form of $x \overset{y}{-} z$, where the number at both ends of the line(i.e. *x* and *z*) are the nodes on the service route and the number above each short line *y* indicates the traffic load on the corresponding link. Notice that some nodes are served locally for the advantage of no latency and zero link failure probability (e.g., $5 \overset{2}{-} 5$). We can see that most of the demands are served by only one rack site. However, the demands at nodes 16 and 22 are split

into two parts and are served by two rack sites since the optimal service provider runs out of capacity.

Node(d)	r_d	demand routing	Node	r_d	demand routing
0	2	$5\overset{2}{-}3\overset{2}{-}0$	16	30	$5\overset{3}{-}4\overset{3}{-}16$ $22\overset{27}{-}17\overset{27}{-}16$
1	4	$5\overset{4}{-}2\overset{4}{-}1$			
2	3	$5\overset{3}{-}2$	17	4	$22\overset{4}{-}17$
3	4	$5\overset{4}{-}3$	18	1	$22\overset{1}{-}17\overset{1}{-}18$
4	4	$13\overset{4}{-}12\overset{4}{-}4$	19	8	$22\overset{8}{-}21\overset{8}{-}20\overset{8}{-}19$
5	2	$5\overset{2}{-}5$	20	1	$22\overset{1}{-}21\overset{1}{-}20$
6	6	$5\overset{6}{-}2\overset{6}{-}6$	21	5	$22\overset{5}{-}21$
7	5	$5\overset{5}{-}7$	22	6	$22\overset{1}{-}22$ $13\overset{5}{-}15\overset{5}{-}26\overset{5}{-}27\overset{5}{-}22$
8	7	$5\overset{7}{-}8$			
9	10	$5\overset{10}{-}14\overset{10}{-}9$	23	6	$13\overset{6}{-}25\overset{6}{-}24\overset{6}{-}23$
10	4	$13\overset{4}{-}12\overset{4}{-}10$	24	2	$13\overset{2}{-}25\overset{2}{-}24$
11	2	$13\overset{2}{-}11$	25	2	$13\overset{2}{-}25$
12	2	$13\overset{2}{-}12$	26	2	$13\overset{2}{-}25\overset{2}{-}26$
13	1	$13\overset{1}{-}13$	27	1	$13\overset{1}{-}25\overset{1}{-}26\overset{1}{-}27$
14	20	$5\overset{20}{-}14$	28	6	$13\overset{6}{-}25\overset{6}{-}26\overset{6}{-}27\overset{6}{-}28$
15	3	$13\overset{3}{-}15$	29	2	$13\overset{2}{-}25\overset{2}{-}26\overset{2}{-}27\overset{2}{-}28\overset{2}{-}29$

Fig.6 Service demands and routing at each node in OBU-SDN.

Input:

All parameters in Section III.A
for each node $d \in \boldsymbol{V}$ in DCN
  Find the nearest data center $v \in \boldsymbol{V}_s$
  if((serve amounts+W_s^d) $\leq \varepsilon$)
    setRouting(d,v, W_s^d)
  else
    $W_s^d \rightarrow$w
    while(w>0)
      distribute the left resource $d_0$of v to node d
      setRouting(d,v,d_0)
      w- $d_0 \rightarrow$w
      find next nearest data center v
      if(there is no enough capacity) continue
      else
        setRouting(d,v, W_s^d)
      end if
    end while
  end if
end for

Output:
the variance of bandwidth utilization of all data centers
the total cost:
transmission cost, waiting delay cost
rout of each node

Conclusion

To improve service QoS and minimize the variance of bandwidth utilization among different data centers with the minimum system cost, it is important to design a scalable and flexible service route. Conventionally, users are served by the nearest data center that is not a better solution because of increasing delay. In this paper, we balance bandwidth utilization among the DCs using SDN controller. We formulated an optimization model to minimize system cost (transmission cost and delay cost) and the variance of bandwidth utilization among different data centers, and then proposed a heuristic service routing algorithm to achieve the same objective, with link failure potential taken into account. Numerical results validated the perfect performance of our proposed algorithms.

References

[1] Roberto Rojas-Cessa, Yagiz Kaymak, and Ziqian Dong,"Schemes for Fast Transmission of Flowsin Data Center Networks", IEEE Communication Surveys & Tutorials, vol.17,No.3, pp.1391-1422, April 2015,

[2] Vendy Wijaya ,"Gartner 2015 Magic Quadrant for Data Center Networking", https://www.gartner.com/doc/reprints?id=1-2F4M3ET&ct=150513&st=sb

[3] Hong Xu and Baochun Li, "Cost efficient datacenter selectionfor cloud services," in *Proceedings of IEEE International Conferenceon Communications in China (ICCC)*, 2012, pp. 51–56.

[4] Bernard Wong and Emin G¨un Sirer, "Closestnode.com: an open access, scalable, shared geocast service for distributed systems," *Operating Systems Review*, vol. 40, no. 1, pp. 62–64, 2006.

[5] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," ACM SIGCOMM Comput. Commun. Rev., vol. 38, no. 4, pp. 63–74, 2008

[6] A. Greenberg *et al.*, "VL2: A scalable and flexible data center network,"*ACM SIGCOMM*, vol. 39, no. 4, pp. 51–62, Oct. 2009.

[7] C. Guo *et al.*, "BCube: A high performance, server-centric network architecture for modular data centers," *ACM SIGCOMM Comput.Commun. Rev.*, vol. 39, no. 4, pp. 63–74, Oct. 2009.

[8] C. Guo *et al.*, "DCell: A scalable and fault-tolerant network structure for data centers," in *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 75–86, Oct. 2008.

[9] J. Hamilton, "An architecture for modular data centers," *arXiv preprint cs/0612110*, 2006.

[10] M. M. Waldrop, "Data center in a box," *Sci. Amer.*, vol. 297, no. 2,pp. 90–93, Aug. 2007

[11] P. Ostovari, A. Khreishah, and J. Wu, "Cache content placement usingtriangular network coding," in Wireless Communications and Networking *Conference (WCNC)*. IEEE, 2013, pp. 1375–1380.

[12] A. Khreishah, I. Khalil, A. Gharaibeh, H. Bany Salameh, and R. Alasem,"Joint caching and routing for greening computer networks with renewable energy source," in *FiCloud*, 2014.

[13] A. Gharaibeh, A. Khreishah, I. Khalil, and J. Wu, "Asymptoticallyoptimal incentive-based en-route caching scheme," in *IEEE MASS*, 2014.

[14] J. Xiao, H. Wen, B. Wu, X. H Jiang, P.H. Ho and L. Zhang, "Joint design on DCN placement and survivable cloud service provision in all-optical mesh networks," IEEE Transactions on Communications, vol. 62, no. 1, pp. 235-245, 2014.

[15] J. Xiao, B. Wu, X. H. Jiang, A. Pattavina, H. Wen, "Scalable Data Center Network Architecture with Distributed Placement of Optical Switches and Racks," *IEEE/OSA Journal of Optical Communcations and Networking*, vol, 6, no, 3, pp. 270-281, 2014

[16] P. Ostovari, J. Wu, and A. Khreishah, "The benefits of cooperation between the cloud and private data centers for multi-rate video streaming," in *Proc. of the he 23rd International Conference on Computer Communications and Networks (ICCCN)*, 2014. *INFOCOM*, 2015.vol. 38, no. 4, pp. 63–74, 2008.

[17] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar,L. Peterson, J. Rexford, S. Shenker, J. Turner, Openflow:Enabling innovation in campus networks, vol. 38, 2008, pp.69–74

[18] Open Networking Foundation, "Software-Defined Networking: The New Norm for Networks,"https://www.opennetworking.org/images/stories/downloads/sdnresources/white-papers/wp-sdn-newnorm.pdf, April 2012.

[19] M. Abu Sharkh, A. Ouda, A. Shami, A resource scheduling model for cloud computing data centers, in: Wireless Communications and Mobile Computing Conference (IWCMC),2013 9th International, 2013, pp. 213–218.

[20] C.Y. Hong, S. Kandula, R. Mahajan, M. Zhang, V. Gill, M. Nanduri,and R. Wattenhofer, "Achieving high utilization with software-driven wan."

[21] JAIN R. Internet 3.0: Ten problems with current Internet architecture and solutions for the next generation[C]// Proceedings of the IEEE MILCOM.Piscataway, NJ, USA. 2006:1-9.

[22] A. Tavakoli, M. Casado, T. Koponen, S. Shenker, Applying nox to the datacenter, in: Proc. of workshop on Hot Topics in Networks (HotNets-VIII), 2009.

[23] Guan Xu,Jun Yang and Bin Dai,"Challenges and Opportunities on Network ResourceManagement in DCN with SDN", 2015 IEEE International Conference on Big Data ,pp. 1785-1790,Oct. 2015

[24] Wang Yong,Tao Xiaoling,He Qian and Kuang Yuwen,"A Dynamic Load Balancing Method of Cloud Center Based on SDN", China Communications,vol. 13,pp. 130-137,Feb. 2016

[25] Renlong Tu ,Xin Wan,Jin Zhao etc.,"Design of a Load-Balancing Middlebox Based on SDN for Data Centers", 2015IEEE INFOCOM WKSHPS,pp.480-485, April 2015

[26] George M. Saridis ,Shuping Peng , Yan Yan and Alejandro Aguado etc."Lightness: A Function-Virtualizable Software Defined Data Center Network With All-Optical Circuit/Packet Switching", Journal of Lightwave Technology ,vol.34 , Issue:7, pp. 1618-1627, April 2016

[27] B. Guo,S. Peng,C. Jackson and Y. Yan etc. ,"SDN enabled Programmable Optical Packet/Circuit Switched Intra Data Centre Network", Optical Fiber Communications Conference and Exhibition (OFC), pp. 1-3, March 2015

[28] W Li, H Qi, K Li, I Stojmenovic, J Lan. "Joint Optimization of Bandwidth for Provider and Delay for User in Software Defined Data Centers". IEEE Transactions on Cloud Computing, 2015.1(1):1-14

[29] Yaofang Li, Jie Xiao, Bin Wu and Hong Wen, "Cloud Service Provision in Two Types of DCN with Awareness on Delay and Link Failure Probability", Photonic Network Communications, Vol. 31, Issue 2, pp. 217-227, Aug. 2015

[30] B. Wu, K. L. Yeung and P.-H. Ho, "ILP formulations for p-cycle design without candidate cycle enumeration," IEEE/ACM Trans. Networking, vol. 18, no. 1, pp. 284-295, Feb. 2010.