A Cumulative Study for Path Isoloation of Network Data

Ankita Shirbhate[1], Sarita Patil[2]

[1]M.E. Computer Engineering, G H Raisoni College of Engg and Management, Pune

[2]Guide and Professor, M.E. Computer Engineering, G H Raisoni College of Engg and Management, Pune

Abstract: ***As the resource competition and conflicts in datacenter networks (DCNs) are frequent and in large extent now a days, this becomes the reason behind the evolution of our concept. On the shared transmission paths, when mixing elephant and mice flows, the different factors taken into consideration are as the latency, throughput and performance degradation. A new flow scheduling scheme is proposed here, Freeway, this advances similar to the path diversity in the DCN topology is carried out simultaneously, like the mice flow completion before the due time and high network utilization of it. Thus this will adjust these two types of paths according to the real-time traffic. The partitioning is done efficiently by Freeway, into low latency and high throughput paths and the different transmission services for each category are provided accordingly. Also thus various flow type-specific scheduling and forwarding methods are used to make full utilization of the corresponding bandwidth.***

Keywords: Data center network, path partitioning, flow scheduling

1. Introduction

In these recent years, Data Center based Internet applications have been fast growing, making its Data Center Network (DCN) operation ever more complex and difficult. Various types of flows are present there as the large and the small named as the Elephant and Mice flows respectively.

Datacenters are called as the core of the cloud computing and the network related to them is an important component to allow applications which are distributed to run efficiently and in predictable way. Though, all the datacenters does not provide cloud computing. So as a matter of fact, there are two important types of datacenters: production and cloud. Production datacenters are sometimes shared by one tenant or among some of the multiple groups, often services and applications, but with comparatively low rate of arrival and departure. The data-analytics jobs are run with relatively small variation in demands and their size fluctuates from about hundreds of servers to the thousands of servers. Cloud datacenters, in contrasting different, have comparatively high rate of tenant arrival and departure, run both the applications that are user-facing and inward computation that requires elasticity as the applications demands are very variable and contains tens to thousands of physical servers. Never the less, clouds can be made up of several datacenters spread around the globe.

There is different type of traffic which is handled by DCN. Eventually, two most common different categories of traffic can be found in DCN[1] are as:
1) Associated with the user tasks, like web surfing or search queries and
2) The machine operations like data backup or MapReduce operations.

The first type of flows, short ones, consists of those generated by user tasks with short duration and should be passed the destination in the assigned Flow Completion Time (FCT)[4], given in the Service Level Agreement (SLA)[2] or deadline with the corresponding users.

The second type of flows, the long flows, which are generated by applications with long duration and mostly require an sufficient throughput, about higher than the minimum acceptable.

The motive is to use the efficient method, like here the Freeway to sort this data separately like for the mice flows to within the specified time limit, i.e. the SLA and with the application of the high throughput for the elephant flows.
And the Freeway is the scheme which is mainly implemented which manages this DCN[1] flows.

Thus is bifurcates the flow of the data to the different pairs as the High Throughput Path (HTP) and Low Latency Path (LLP). These are thus used as per the real time traffic load. The less frequent are the Elephant flows that are scheduled centrally while the ECMP[6] algorithm is used to schedule the mice flows.

But there are some challenges to deal with these mice flows as well, such as the defined SLAs by which the flows should reach to the destination. This can be achieved by reducing the concurrent flows and formulate a queuing model to get the highest delay threshold. Also the challenges are there like to deal with the Low Latency Paths to the various traffic loads thus present and based on the commodity switches, partitioning on the paths as well.

The transmission delay bounds are highly responsible thus to the mice flows due to this Service Level Agreement.

This paper is segmented as follows:
Section 2 dedicates literature survey.
Section 3 is for the proposed methodology analysis.
Section 4 gives us the conclusion of the paper.

2. Literature Survey

There are various concepts used in this paper.

Previously, as default routing algorithm, Equal Cost Multi Path (ECMP) is used in data centers, where to shortest paths, the pair of servers is routed. With this ECMP[6], two long flows can be routed over the similar path causing hot-spots in the network. As a result, the throughput is decreased and the corresponding path latency which is traversing through the congested link is increased, increasing the FCT as well.

This usually thus impacts the user experience when the small flows related to the user tasks are routed to the congested paths.
Previous work on this has three ways of solutions:
1) Sacrifice among the two constraints to optimize one of them,
2) Increase the complex traffic engineering with priority scheduling, and
3) Modify the present DCN architectures like the TCP stack, OpenFlow switches etc.

Also there are hybrid types of the solutions for mixed short and long flows in the DCNs known as DiffFlow[10], that are used to achieve the desired relation between the low latency and high throughput. In the proposed approach, this use of the traditional ECMP for short flows, which minimises the FCT with a very small amount of out-of-order problem generally encountered in every packet basis approaches, while for long flows, the Random Packet Switching (RPS)[5] is used to balance the load without affecting the path latencies of those of the short flows.

To get this implemented the application on the OpenFlow[12] switches of the packet sampling technique to trace the long flows and the SDN controller to spread the awareness on this presence of long flows and put the RPS[5] rules on the switches. This can be thus shown with the analytics and the simulations that the proposed method can effectively balance the load of the network, despite of keeping FCT and the corresponding throughput within the pre-defined ranges itself.

Like the Hedera[8] which is used as dynamic flow scheduling for data center networks. It efficiently schedules a multi-stage switching fabric making productive use of the available network resources, thus being scalable and the dynamic flow scheduling system. But Freeway effectively reduces the delay of mice flow efficiently and achieves much higher throughput compared with Hedera, which is compared with the help of our simulation results.

Routing Bridges also known as the (RBridges)[3] provide efficient pair-wise forwarding without any configuration and the safe forwarding even for the periods of temporary loops and to support the multipaths of both unicast and multicast traffic. They aquire these goals using IS-IS routing and encapsulation of traffic with a header that contains a hop count. These RBridges are compatible with old IEEE 802.1 customer bridges with the IPv4 and IPv6 routers and end nodes. They are as not visible to the current IP routers as similar to bridges and similar to routers, the bridge spanning tree protocol is terminated by them.

This design supports VLANs and the optimization of the distributed multi-destination frames based on VLAN ID and which are based on IP-derived multicast groups. It also allows the unicast forwarding tables at those transit RBridges to be sized as per to the number of RBridges[3] (instead of the number of end nodes), which allows their forwarding tables to be comparatively smaller than in conventional customer bridges.

The another concept used thus is the existing transport layer solutions proposed to deal with the problems of TCP in data center networks contains as about the issues of TCP in data center networks; to introduce different transport layer solutions and the comparison to discuss the challenges of existing solutions proposed to improve the performance of TCP in data center networks.

The traffic characteristics are changed by thus TinyFlow[12] of data center networks to be done by ECMP by breaking elephants into mice. Network that contains the large number of mice flows only, ECMP hence balances the load and performance is enhanced thus.

The bandwidth and Network latency are mostly the important concerns to be considered for the goodness of network. The HULL[4][7] (High-bandwidth Ultra-Low Latency) architecture is thus introduced to balance two seemingly different motives as to: high bandwidth utilization and the fabric latency. The Phantom Queues which transmits the congestion signals before utilization of the network links and which forms queues at the switches while keeping the 'bandwidth headroom' aside is the feature of HULL.

The latency sensitive traffic that avoids buffering and the corresponding large delays are achieved with utilizing near about less than link capacity. Also the recent congestion control algorithm is implemented which responds to congestion and to control the bandwidth penalties which that are caused due to the buffer less functioning.

To deal with the Online Data-Intensive (OLDI) applications, this generally implements the tree-based algorithms that results in bursts of traffic with due to the deadlines. Recently this is worked upon either deadline-agnostic (DCTCP)[11][14] or deadline-aware (D3) but has to suffer under bursts due to different race conditions. To add more, D3 has some practical drawbacks of additional required changes to the switch hardware and to be unable to coexist with legacy TCP.

Deadline-Aware Datacenter TCP (D2TCP), is a new transport protocol, which is capable of handling the bursts, is a deadline-aware protocol, and is also readily deployable. D2TCP uses a distributed approach to allocate the bandwidth which basically enables D2TCP's properties. Also D2TCP employs a new algorithm which avoids the congestion, that makes the use of ECN feedback and there are deadlines which modulates the congestion window through a gamma-correction function. When used for a small-scale implementation and at-scale simulations, this D2TCP reduces the fraction of missed deadlines in comparison with DCTCP and D3 by 75% and 50%, respectively.

The various papers compared in this paper are as:
[1] Wei Wang, Yi Sun, Kavé Salamatian, and Zhongcheng Li, Member, IEEE, "Adaptive Path Isolation for Elephant and Mice Flows by Exploiting Path Diversity in Datacenters",IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, VOL. 13, NO. 1, MARCH 2016. In this paper, high data congestion due to traffic over flow of network is overcome by Freeway method.

With the Advantages:
1) Bifurcates flow into elephant & mice flows.
2) Efficiently manages flows as per low latency & high throughput.

Disadvantages:
1. Eventually latency & throughput are parameters can be improved. upon thus overall system performance.

[2] Y. Cao, M. Xu, X. Fu, and E. Dong, "Explicit multipath congestion control for data center networks," in Proc. 9th ACM Conf. Emerging Netw. Exp. Technol. (CoNEXT), 2013, pp. 73–84. In this paper, issue is with the existing data transfer proposals can't utilize path diversity of DCNs in all extent to enhance the throughput of large flows or small flows & their low latency requirement. And the methods implemented are as BOS algorithm & TraSh algorithm.

With the Advantages:
1) Balance throughput with latency
2) Implemented as the congestion control scheme of MPTCP.

Disadvantages:
1) Extensive experiments conducted demonstrate the performance of XMP.
2) Parameter setting has impact on the performance of XMP.

[3] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, "Hedera: Dynamic flow scheduling for data center networks," in Proc. Netw. Syst. Des. Implement. (NSDI), 2010, vol. 10, p. 19. In this paper, the issue is with existing IP multipath protocols cause bandwidth losses due to long term collisions. And it thus efficiently schedules a multi-stage switching fabric to utilize network resources & implemented using commodity switches and unmodified hosts.

With the Advantages:
1) Efficiently utilize snetwork resources .
2) Better than other static load-balancing methods

Disadvantages:
1) Performance gain relies upon the rates and durations of the flows in the network.
2) Gives efficient results when the network is stressed with many huge data transfers.

[4] M. Alizadeh et al., "CONGA: Distributed congestion-aware load balancing for datacenters," in Proc. SIGCOMM, 2014, pp. 503–514. In this paper, ECMP balances load poorly whereas centralized schemes are very slow for the traffic volatility in datacenters. Uses Clos topologies and overlays for network virtualization.

With the Advantages:
1) Efficiently balances load.
2) Seamlessly handle asymmetry even without necessity of any TCP modifications.

Disadvantages:
1. This requires global congestion-awareness to handle asymmetry.

[5] M. Alizadeh, A. Kabbani, and T. Edsall, "Less is more: Trading a little bandwidth for ultra-low latency in the data center," in Proc. Netw. Syst. Des. Implement., 2012, p. 19. In this paper, Network latency should be a concern in networking and high bandwidth utilization. Phantom Queues delivers congestion signals before network links are fully utilized, also DCTCP is used which responds to congestion.

With the Advantages:
1) Delivers baseline fabric latency and high bandwidth utilization.
2) Supports high fabric goodput for bandwidth-sensitive applications.

Disadvantages:
1) Requires to be tested on a larger multi-switch testbed and with more diverse workloads.
2) Needs to quantify the buffering requirements for communication patterns.

[6] Francisco Carpio, Anna Engelmann and Admela Jukan, "DiffFlow: Differentiating Short and Long Flows for Load Balancing in Data Center Networks," in Proc. SIGCOMM, 2016. In this paper, ECMP does not differentiate between short and long flows, this issue turns to cause hot-spots in the network. This makes use of SDN and packet sampling technologies, thus can detect and forward long flows while short flows are forwarded using ECMP.

With the Advantages:
1) Improves the FCT for short flows and throughput for long flows.
2) Receive Package Steering used for long flows can efficiently help to load balancing the entire network.

Disadvantages:
1) Necessary to use a centralized controller for the advertisement of long flows.
2) Causing this solution cannot be applied in current DCNs.

3. System Architecture

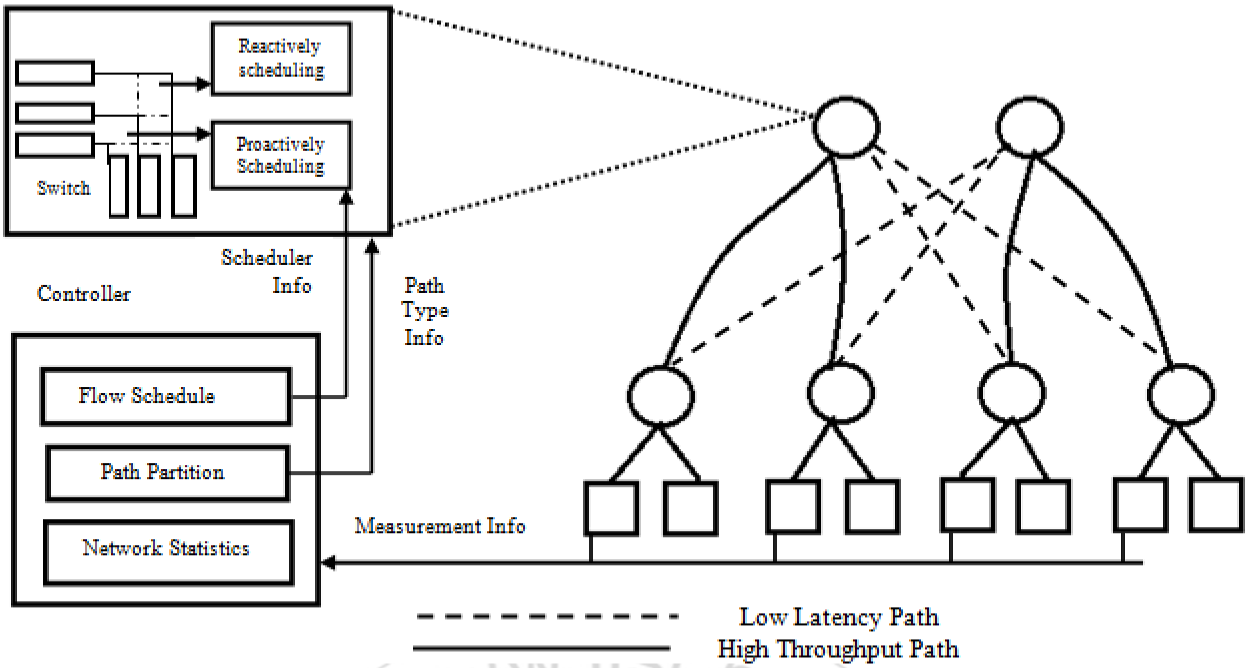

Fig. Freeway Architecture. Shown are the network wide statistics and real time traffic load, dynamically partitioned pats to Low Latency Path and High Throughput Path.

4. Conclusion

Due to drastic increase in digital devices there is been urge to store the data at remote location like never before. So data centers are the boon to store the same by adopting many revolutionary techniques like cloud computing, distributed computing and even green computing. This paper analyses the requirement of data's path isolation based on the size of the network data by studying various work of the prior authors. . After analyzing most of the systems, this paper comes to a conclusion that much progression need to be achieve to attain perfection, So as a conclusion and as effort of contribution this paper thinks on path isolation based on decision tree classification and genetic algorithms in the coming editions.

References

[1] Wei Wang, Yi Sun, Kavé Salamatian, and Zhongcheng Li, Member, IEEE, "Adaptive Path Isolation for Elephant and Mice Flows by Exploiting Path Diversity in
[2] Datacenters",IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, VOL. 13, NO. 1, MARCH 2016.
[3] Service Level Agreement for VMware Vcloud Air, VMWare, 2015 [Online]. Available:http://www.vmware.com/cn/support/vcloud-hybridservice/sla
[4] Routing bridges: Base protocol specification," IETF, RFC 6325, 2011
[5] M. Alizadeh, A. Kabbani, and T. Edsall, "Less is more: Trading a little bandwidth for ultra-low latency in the data center," in Proc. Netw. Syst. Des. Implement., 2012, p. 19.
[6] G. Appenzeller, I. Keslassy, and N. McKeown, "Sizing router buffers," in Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun. (SIGCOMM), 2004, pp. 281–292
[7] Y. Cao, M. Xu, X. Fu, and E. Dong, "Explicit multipath congestion control for data center networks," in Proc. 9th ACM Conf. Emerging Netw. Exp. Technol. (CoNEXT), 2013, pp. 73–84.
[8] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in Proc. SIGCOMM, 2008, pp. 63–74.
[9] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, "Hedera: Dynamic flow scheduling for data center networks," in Proc. Netw. Syst. Des. Implement. (NSDI), 2010, vol. 10, p. 19.
[10] M. Alizadeh et al., "CONGA: Distributed congestion-aware load balancing for datacenters," in Proc. SIGCOMM, 2014, pp. 503–514.
[11] Francisco Carpio, Anna Engelmann and Admela Jukan, "DiffFlow: Differentiating Short and Long Flows for Load Balancing in Data Center Networks," in Proc. SIGCOMM, 2016.
[12] M. Alizadeh, A. Greenberg, and D. A. Maltz, "Data center TCP (DCTCP)," in Proc. SIGCOMM, 2011, vol. 41, pp. 63–74.
[13] H. Xu and B. Li, "TinyFlow : Breaking Elephants Down Into Mice in Data Center Networks," in in Proc. IEEE LANMAN, 2014.
[14] H. Xu and B. Li, "RepFlow: Minimizing flow completion times with replicated flows in data centers," Proceedings - IEEE INFOCOM, 2014.
[15] B. Vamanan, J. Hasan, and T. Vijaykumar, "Deadline-aware datacenter TCP (D2TCP)," in Proc. SIGCOMM, 2012, pp. 115–126.