International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I3P207

Volume 69 Issue 3, 34-38, March 2021
© 2021 Seventh Sense Research Group®

A ShuffleNet based Hybrid Architecture for Data Center Network

Tapasmini Sahoo[#1], Bibhu Prasad Mohanty[*2]

[#] *Assistant Professor, Electronics and Communication Engineering Department, SOA (Deemed to Be) University, Odisha, India*

[*] *Professor, Electronics and Communication Engineering Department, SOA (Deemed to Be) University, Odisha, India*

[1]tapasminisahoo@soa.ac.in, [2]bibhumohanty@soa.ac.in

Abstract — *This paper describes an effective ShuffleNet-based architecture of an electro-optic data center. Usually, the data center network (DCN) architecture needs to provide high bandwidth, high speed to meet high internet service demand and numerous web-based applications. Thousands or millions of servers often had to support the architecture. The architecture of ShuffleNet has the unique benefit of shuffling the lightpath that impressively reduces the cost of computing with a high degree of accuracy. The suggested architecture with ShuffleNets based on the WDM technique ensures better throughput and packet loss performance with low latency and power consumption. The architecture also includes a layer of fat-tree which improves the network's scalability.*

Keywords — *ShuffleNet, Electro-Optic, Lightpath, WDM.*

I. INTRODUCTION

Currently, cloud-centric data centers (DCs) perform various information technology-related services. Data centers form the backbone of an extensive variety of Internet applications like Web hosting, social- networking, e-commerce, and various grid or cloud computing-related services. To meet this demand, the size and complexity of the data centers are increasing in a rigorous manner. So it is essential to understand the existing issues and upcoming shortfalls, and challenges for designing the data centers.

In general, the DCN architectures have been proposed in two broad categories, viz., switch-centric architecture and server-centric architecture. Fat–Tree is a classic switch-centric hierarchical topology using identical commodity switches at all levels (edge, aggregate, and core) for full bisection bandwidth, however with huge wiring complexity when scaled up [3]. VL2 is another switch-centric architecture, using commodity switches to form a three-layered tree topology offering a complete bipartite graph between core and aggregate switches [4]. To increase the fault tolerance of DCNs, Aspen Tree has been proposed with hierarchical topology with in-built fault tolerance, however, at the cost of scalability of the network [5]. Some of the candidates, server-centric DCN architectures include D-Cell, BCube, and several others [1]. D-Cell is a server-centric hierarchical topology employing fewer switches along with servers having network interface cards (NICs) as ports, wherein the topology is constructed through a recursive scheme offering excellent scalability [6]. B-Cube is another recursively-constructed topology, which uses a few mini-switches along with the servers having multiple NICs [7].

Due to the unprecedented growth of cloud-centric applications, the next-generation DCNs would require low latency and high capacity (speed) along with a scalable architecture. So far, the DCNs have been designed with electrical packet-switching, but the interconnections between the servers, switch, and between switches used optical fiber links. Given the fact that the DCNs should accommodate a huge number of servers, such architectures cannot be recommended as a scalable network for future growth, as the network complexity with electrical-switching equipment turns out to be a serious issue due to limited bandwidth in electrical switches, high power consumption, and wiring complexity. On the other hand, optical switching technology offers much higher capacity, lower cost, and power consumption. However, the optical switches, typically using micro-electro-mechanical switches (MEMS), suffer from high latency (10 ms) at the time of switch reconfiguration and hence cannot handle bursty traffic efficiently.

C-through offered an evolutionary work in the category of electro-optic or hybrid DCNs using optical as well as electrical switching [2]. Helios is another hybrid architecture using a two-level multi-rooted tree topology with pod and core switches [8]. Some futuristic topologies have also been examined in the literature, viz., OSA, Mordia, LION, etc. [1], all of them employing fully-optical switching architecture promising extremely high speed, while one is not sure at this stage how far these architectures can be scaled in the optical domain itself with the evolving DCN demands. In the foreseeable future, it is therefore conjectured that the DCNs need to grow with hybrid architectures to enhance the speed and size while keeping the power consumption within the affordable limit.

ShuffleNet [9, 10] is a well-known multi-hop virtual topology that uses Wavelength Division Multiplexing (WDM) [11, 12] with intensity modulation as the underlying physical topology. A basic ShuffleNet is designated as (p, k) ShuffleNet consisting of (k.pk) number of nodes. They are arranged as the k number of columns and the pk number of nodes in each column, and the kth column is wrapped around to the first in a cylindrical way [13]. This architecture can overcome both wavelength-agility and pre-transmission-coordination problems.

This paper depicts a novel hybrid architecture of DCN based on the hierarchical use of ShuffleNets. The proposed architecture offers high scalability with appreciable bandwidth by virtue of fat trees. The use of ShuffleNet in the architecture has the unique advantage of shuffling the lightpath that greatly reduces computational cost with a high level of accuracy. The proposed architecture also ensures low power consumption due to the substantial use of optical devices.

II. PROPOSED SHUFFLENET BASED HYBRID ARCHITECTURE OF DCN

The architectures of existing data centers suffer from some disadvantages such as the end-to-end delay, bandwidth, energy consumption, and quick failover [7]. Large organizations with voluminous traffic are required to interconnect their own data centers, which are discretely stationed in remote locations, to maintain system efficiency.

The provision of multiple channels hopping through time-sharing mode in the existing shuffle net architecture has provided some respite in traffic management in such organizations because of its wavelength-agility and better transmission coordination. But, in turn, this architecture suffers a significant end-to-end delay. In the following, Figure 1, we have depicted our proposed architecture based on ShuffleNet and augmented with fat-tree, which, additionally, reduces the end-to-end delay incurred in ShuffleNet.

Due to the emerging demand, servers now require a low delay and high bandwidth communication. Optical connectivity consumes less power at the same bandwidth provided by it is larger than electrical links. However, the optical switches, typically using micro-electro-mechanical switches (MEMS), suffer from high delay (~10 ms) at the time of switch reconfiguration and hence cannot handle bursty traffic efficiently. In our architecture, we propose to bypass the traffic in such a way that end-to-end delay will not be much suffered by MEMS delay and better load balancing is achieved.

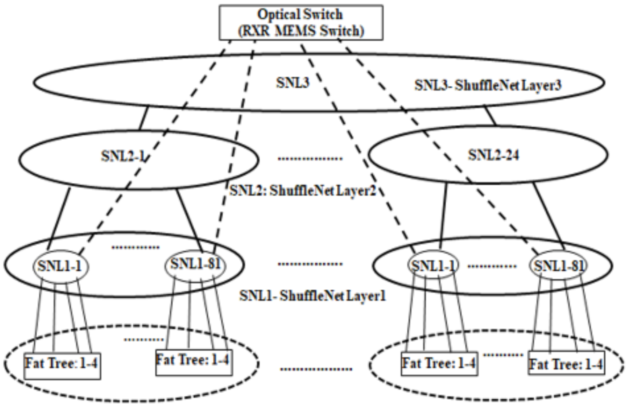

Fig. 1. Structure of the Proposed Hybrid DCN Architecture.

The proposed architecture consists of three discrete portions: 1st portion consists of fat trees with several numbers of top of rack (ToR) switches, and each ToR switch handles numbers of servers. All servers are connected to the ToR through optical links. Fat trees are electrical switching enabled sub-network, and the ToRs connected to this subnetwork support packet switching. The next 2nd portion is the hierarchical ShuffleNet based optical network. ToR switches are connected to the hierarchical ShuffleNet based optical network in parallel with a fat tree-based electrical switching enabled sub-network. The integration of these two architectures is used to handle two different types of traffic present in the network. The traffics are classified into two types: small size bursty traffic, commonly called mouse traffic, and the large volume of traffic called elephant traffic. All bursty traffic follow packet switch enabled fat tree-based electric switch domain. And all large volumes of traffic follow the ShuffleNet based optical network. This type of traffic segregation and transmission significantly enhance the switching speed and reduce the power consumption of the network.

The proposed architecture, as shown in Fig 1, consists of fat trees at the lower end with k=8. Four such fat trees are connected to the next 1st ShuffleNet in the 2nd layer. There are 81 ShuffleNets in the 2nd layer with the configuration as k=3 and on the strength of traffic. If the traffic is elephant traffic, then we have to forward it to MEMS, and through MEMS only, it will be further connected to another ShuffleNet to ToR. We can't avoid this increase in the delay that is offered by the MEMS switch. But, if the traffic is mouse traffic, then it will be advised not to direct the flow to MEMS rather than directed through another ShuffleNet such that we can avoid an extra increase of delay. Within the ShuffleNet, it is better to have so that the end-to-end delay will be reduced, so for this purpose, the 3rd layer of ShuffleNet is highly required.

III. OPERATION OF THE PROPOSED ARCHITECTURE

In the proposed architecture, the "n" number of ToR switches are used, which are associated with the fat tree. Each ToR switch can support the "m" number of servers. Each ToR support an equal number of users. Each ToRs are connected in parallel with both the electric fat tree enabled domain and ShuffleNet based optical domain.

Fig.2 shows the hierarchical ShuffleNet structure for the proposed hybrid architecture. Here there is "x" number of ShuffleNets are used to support the "n" number of ToRs in the first layer. They are designated as SNL1-1, SNL1-2 to SNL1-x. The second layer of ShuffleNets is designated as SNL2-1 to SNL2-y. All these ShuffleNets are controlled by the third layer ShuffleNet designated as SNL3.

When the data is transferred from one ToR to another between two different ShuffleNet, packets are transferred through ShuffleNet 3. An optical switch is also connected in parallel with the third ShuffleNet to reduce the load in the third ShuffleNet. If the packet size is too large, then it follows the optical switch.

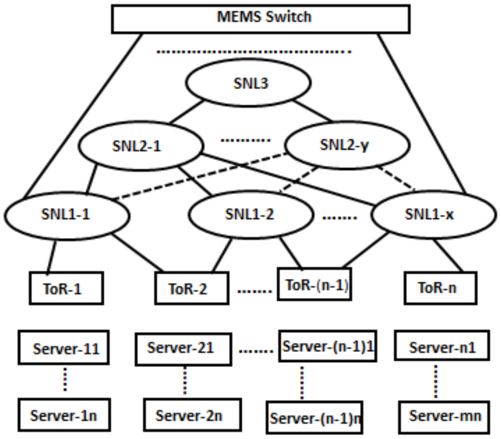

Fig. 2. Hierarchical ShuffleNet Structure for the Proposed DCN Architecture.

In the proposed framework for end-to-end delivery of the packet, the following steps are followed:

a) Traffic monitoring and management

The network estimates the demand of end-user traffic in an application transparent manner by increasing the buffer limit per connection socket and the occupancy time per connection socket. The per-flow basis queuing technique used in the proposed model has the advantage of preventing blocking between concurrent flows. Therefore, low bandwidth latency-sensitive data is not at all experience any extra delay due to high bandwidth data flow.

b) Traffic demultiplexing

There are three paths on which a packet can travel in the optical domain. Firstly, the packet is transferred from one ToR to another under the same ShuffleNet, secondly from one ToR to another under different ShuffleNet, and the packet is transferred through the third ShuffleNet.

Depending upon the request and type of traffic from the servers, each ToR assigns the path for the traffic. If the traffic is bursty in nature and latency-sensitive, ToR assigns the electric ports for the traffic similarly for high bandwidth large volume of traffic ToR assign the optical port for the traffic transfer.

c) Path Selection

In the proposed architecture, when there is a request from any server to the respective ToR, first, it will check the destination address of the server. If the address of the destination server belongs to the same ShuffleNet, it will forward the traffic to the ShuffleNet node, and the packet is reached to its destination through the ShuffleNet routing algorithm. If ToR finds that the destination address does not belong to the same ShuffleNet, it will forward the traffic to the third ShuffleNet. Fig. 3 shows the ShuffleNet architecture for the proposed model.

To understand the process of the packet transfer, let us consider a simple example that server A wants to send some data to server B through the optical domain. Server A is connected to ToR1, and server B is connected to ToR 3 and lets both the ToRs are under the same ShuffleNet. Then for the transmission of packets, server A sends a request to ToR1. ToR1 checks the destination address. If ToR1 finds that the destination ToR is under the same ShuffleNet, it forwards the packet to the corresponding ShuffleNet node. After proper path selection, the data reached destination server B through ToR3.

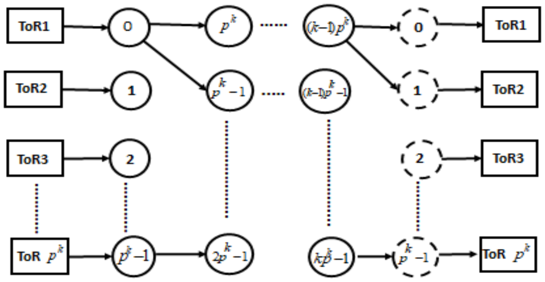

Fig. 3. ShuffleNet configuration for the proposed architecture.

d) Packet transfer

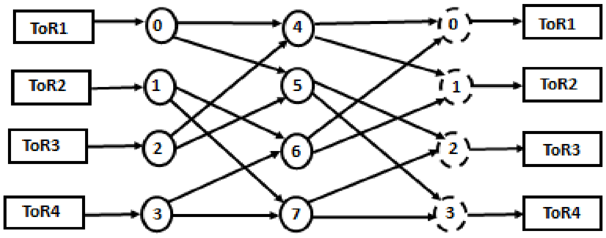

Fig. 4. (2,2) ShuffleNet structure.

For simplicity, consider an example that Server A wants to send some data to Server B through Optical Domain. Server A is connected to ToR1, and Server B is connected to ToR 3. Both ToRs are under the same ShuffleNet. The structure of the ShuffleNet is (2, 2). The connection of ShuffleNet and ToRs for (2, 2) ShuffleNet is shown in Fig.4.

For sending the packet, Server A sends a request to ToR1. ToR1 checks the destination address. If ToR1 finds that the destination ToR is under the same ShuffleNet, it forwards the packet to the corresponding ShuffleNet node. The signal flow graph for this particular case of packet transfer is described in Fig.5.

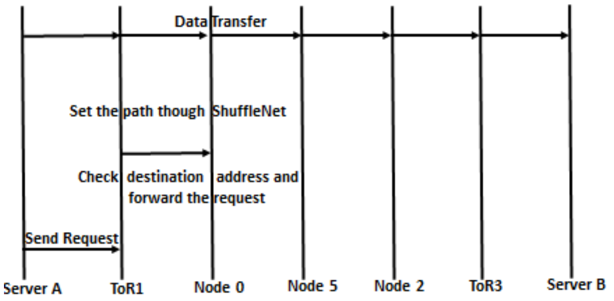
Data Transfer
Set the path though ShuffleNet
Check destination address and forward the request
Send Request
Server A
ToR1
Node 0
Node 5
Node 2
ToR3
Server B

Fig. 5. Signal flow diagram for path set up and data transfer.

e) Packet transfer between different shufflenet

In this case, some ports of the ShuffleNet nodes are connected to the higher hierarchical ShuffleNet. So when communication is required from one ShuffleNet to another, the packet is forwarded through higher-level ShuffleNet. Let us consider a situation, for example, when Server A in ShuffleNet 1 connected to ToR12 wants to communicate with Server B, which is connected to ToR24 in ShuffleNet 2. The path of packet transfer from ToR12 to ToR24 through higher-level ShuffleNet is briefly shown in Fig 6. The thick solid lines show the selected path for data transmission between the two servers.

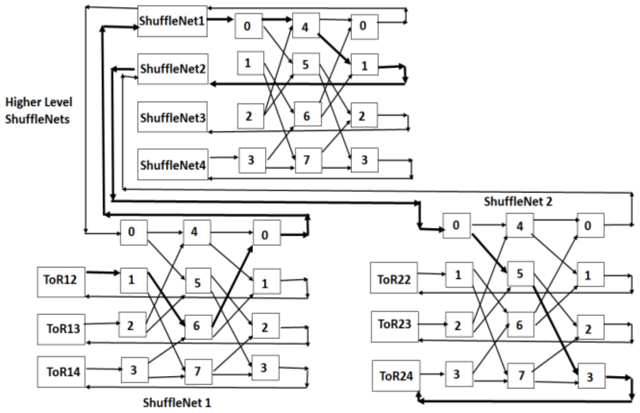

Fig. 6. Packet transfer between two different ShuffleNet transfers.

IV. PERFORMANCE EVALUATION

To explore the performance of the proposed ShuffleNet based hybrid architecture, the analysis of packet loss and throughput of the architecture is considered as a function of the traffic patterns and the number of servers.

The traffic is divided into three groups in the simulation as inter-ShuffleNet, intra-ShuffleNet (mainly for layer 1 ShuffleNet in conjunction with fat trees), and intra-ToR. The packet destinations are selected at random in each group. Since most of the traffic is exchanged within the ToR and ShuffleNets, in Table I, various traffic ratios are considered as shown below.

The performance of the proposed DCN, consisting of 2500 TORs interconnected by layer 1 of ShuffleNets with (p=3,k=3) under different traffic patterns, is investigated. Each ToR connects 40 servers with 10 Gb/s links, resulting in a DCN composed of 100000 servers.

TABLE I
TRAFFIC PATTERNS

Traffic			**Case 1**	**Case 2**	**Case 3**	**Case 4**
Inter ShuffleNet			15%	10%	15%	10%
Intra ShuffleNet	35%	40%	25%	30%		
Intra ToR	50%	50%	60%	60%		

If the total amount of inter-ShuffleNet and intra-ShuffleNet traffic increases, the packet loss increases. The packet loss, as shown in Fig. 7 in case 1 and case 2 (50 percent traffic out of ToR), is also greater than in case 3 and in case 4 (40 percent traffic out of ToR). Although the amount of traffic coming out of the ToR is the same for cases 1 and 2, the packet loss in case 1 is significantly greater than in case 2.

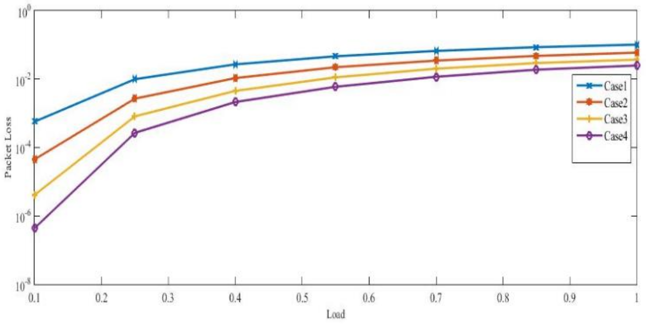

Fig. 7. Packet loss of the proposed architecture under different traffic patterns.

The inter-ShuffleNet traffic in case 1 is greater than in case 2, as most of the inter-ShuffleNet traffic has more link capacity. Thus, the real load of the system in case 1 is higher than the case 2. Similarly, the same analysis can also be applied to the performance results under case 3 and case 4 traffic patterns.

Figure 8 illustrates the network throughput as a function of the traffic load and DCN size. The throughput saturates at the load of 0.6, and the performance is better in highly scalable DCN, i.e., with higher DCN size by considering more number of servers in the network.

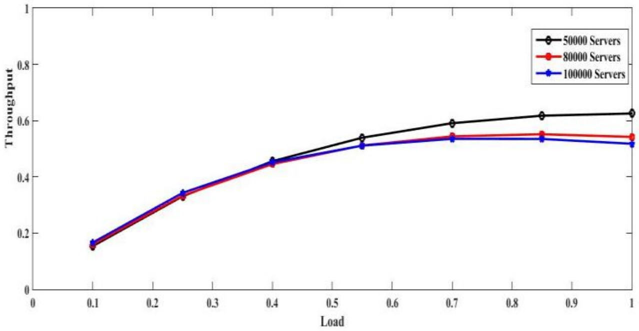

Fig. 8. Normalized throughput of the proposed architecture under different DC sizes.

V. CONCLUSION

The use of the optical switching technique is the best approach for consistent extension of transmission capacity in DCs. The hybrid architecture effectively combines the benefits of both electrical and optical switching technologies and is used as a fully equipped packet-switched network with the same performance measures at low cost, complexity with lower power consumption, and a lower rate of packet loss. This defines the traffic subset ideally suited for the switching of circuits and dynamically reconfigures the network topology based on evolving patterns at runtime. It does not require any modification of the end host, and rather it requires modification of the switch software. From the scalability point of view ShuffleNet, based architecture can support more nodes than that of optical switching DCN model. So it can be said that the ShuffleNet based architecture can be a feasible solution for the next generation DCN architecture.

REFERENCES

[1] Wenfeng Xia, Peng Zhao, Yonggang Wen, A Survey on Data Center Networking (DCN): Infrastructure and Operations IEEE Communications Surveys & Tutorials 19(1)(2017).

[2] Guohui Wang, David G. Andersen, Michael Kaminsky, Konstantina Papagiannaki, T. S. Eugene Ng, Michael Kozuch, Michael Ryan, c-Through: Part-time Optics in Data Centers, in Proc ACM SIGCOM Conf New Delhi India,(2010) 327-38.

[3] M.Al Flair, A. Loukissas and A Vahdat, A Scalable commodity data center network architecture, in Proc ACM SIGCOM, Conf. Data Communication, Seatle, WA, USA, (2008) 63-74.

[4] Albert Greenberg, James R. Hamilton, Navendu Jain, Srikanth Kandula, Changhoon Kim, Parantap Lahiri, David A. Maltz, Parveen Patel, and Sudipta Sengupta, VL2: A Scalable and Flexible Data Center Network in Proc ACM SIGCOM Conf Data Communication (SIGCOMM)Barcelona, Spain (2009) 51-62.

[5] Meg Walraed-Sullivan, Amin Vahdat, Keith Marzullo, Aspen Trees: Balancing Data Center Fault Tolerance, Scalability and Cost' in proc 9th ACM Conf. Emerg. Netw.Exp Technology (CoNEXT), Santa Barbara, USA, (2013) 85-96.

[6] Chuanxiong Guo, Haitao Wu, Kun Tan, Lei Shi Yongguang Zhang, Songwu Lu, 'DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers', in Proc ACM SIGCOM, Conf. Data Communication ,Seatle, WA, USA,(2008) 75-86.

[7] C. Guo et al., BCube: A high performance, server-centric network architecture for modular data centers, in Proc. ACM SIGCOMM Conf. Data Commun. (SIGCOMM), Barcelona, Spain, (2009) 63–74.

[8] N. Farrington et al., Helios: A hybrid electrical/optical switch architecture for modular data centers, in Proc. ACM SIGCOMM Conf., New Delhi, India, (2010) 339–350.

[9] K. Sivarqan and R. Rarnaswami. 'Multihop Lightwave Networks Based on d e Bruijn Graphs., submitted to Eb .on C ommun.

[10] M. G. Hluchyi, M.J. Karol, ShuffleNet: An Application of Generalized Perfect Shuffles to Multihop Lightwave Technology", IEEE (1991)379-390.

[11] B. Mukhejee., Architectures and Protocols for WDM-Based Local Lightwave Networks, Part 11: Multihop Systems," [tentative title). to appear, IEEE Network. (1992).

[12] N. R. Donoetal., A Wavelength Division Multiple Access Network for Computer Communication,lEEEI. Sel. Areas in Commun. 8(1990) 983.994.

[13] K S. Acampora and M. I. Karol, An Overview of Lighiwave Packet Networks,IEEE Nehvork Mag., 3(1)(1989) 29-41.

[14] B. Ramamurthy, D.Datta, H. Feng, B. Mukherjee, Impact of Transmission Impairments on the Teletraffic Performance of Wavelength –Routed Optical Network, IEEE Journal of Lightwave Technology, 17(1999) 1713-1723.