International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 11
July 2016

A Investigation and Evaluation of Data Center Network architectures

[1] Dr. K. Praveen Kumar, [2] K. Sree Ranganayaki
[1]Associate Professor, School of Electrical Engineering & Computing, Department of Computing, Adama Science &Technology University, Adama, Ethiopia

[2] M.Tech Student, Chaitanya Institute of Technology and Science, Warangal, India

ABSTRACT: Large-scale data centers shape the center infrastructure support for the ever expanding cloud based totally services. Thus the overall performance and dependability characteristics of data facilities will have great impact at the scalability of those offerings. In specific, the data center networks wishes to be agile and reconfigurable that allows you to reply fast to ever changing software needs and service necessities. Significant studies portraits has been achieved on designing the data center network topologies that allows us to enhance the performance of data centers. In this paper we provide an explanation for a regular data center network architecture within the industry, the challenges current data center networks stumble upon today and introduce proposed solutions by means of a recent research.

KEYWORDS- Data Center Network, Topology, Fault Tolerance, cloud computing, optimization

I. INTRODUCTION

Data centers are large facilities hosting massive numbers of servers and the associated support infrastructure. Servers can be used for several purposes, including interactive computation, batch computation, and real-time transaction. Data centers can be seen as a composition of information technology (IT) systems and the support infrastructure. The IT systems provide services to the end users and the support infrastructure supports the IT systems by supplying power and cooling. The IT systems include servers, storage devices, networking devices, middleware and software stacks, e.g., hypervisors, operating systems, and applications. The support infrastructure includes backup power generators, uninterruptible power supplies (UPSs), power distribution units (PDUs), batteries and the cooling technology (CT) systems. In general, the goal of the data center operators is the maximization of the data center efficiency, subject to a set of operational constrains and to constraints on the reliability and on the availability of the provided services. The large number of constraints and their heterogeneity in nature, e.g., some constraints may be related to physical constraints of the data center subsystems and others may be related to software (cyber) constraints, make data center control a cumbersome task and a challenging research problem. We believe that control algorithms based on cyber-physical models of data centers can improve current control strategies by leveraging information about how the workload distribution affects the overall power consumption and the quality of service (QoS) provided to the users.

In this paper, we present history and taxonomy of various DCN topologies that have been proposed so far and how they have advanced the state-of-the-art technology to overcome aforementioned challenges. The main focus while designing the DCN architecture, has been scalability, cost, latency, extensiblity.

II. RELATED WORKS

However, despite the fact that maximum data facilities use Ethernet switches to interconnect the servers, there are still many exclusive ways to put in force the interconnections, leading to specific data center network topologies. Each of these different-topologies is characterized by means of different aid necessities, aiming to bring enhancements to the performance of records facilities. In the subsequent sections a few representative topologies used in data facilities will be discussed. If servers and switches are appeared as vertices, wires as edges, the topology

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 11
July 2016

of every data center network may be represented as a graph.

Figure 1 gives a taxonomy of the different data center network topologies. Furthermore, here we summarise the notations used in this section:

- n: The number of ports in a switch in an architecture.
- k: The number of ports in a server in an architecture.
- N: The total number of servers inside a data center network.

It should be noted that n and k may vary according to the position of the node.

ack.

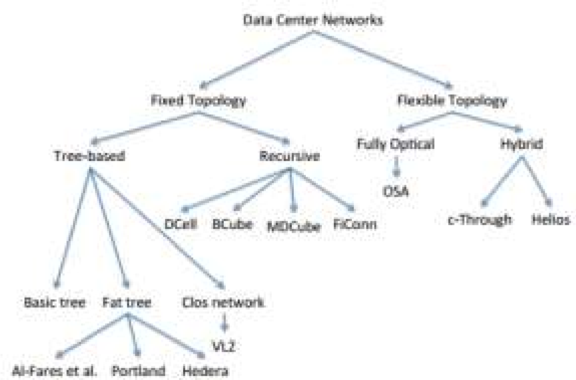

Fig. 1. A Taxonomy of Data Center Topologies

Ethernet: Ethernet is a facts hyperlink layer protocol to send packets (frames) from one factor to every other point (host or switch), without delay connected to every different. It affords excellent attempt provider primarily based on collision detection (CD) without any flow manipulate. Frames are dropped if the queues are complete at the receiver without notifying the sender. Further, Ethernet switches are much like a NoC crossbar and put in force FIFO model for packet processing.

TCP/IP: IP is a network layer protocol to ensure routing of packets from one host to some other host within the network. TCP runs on pinnacle of IP layer and implements: (a) drift control to save you receiver's buffer from overflowing, (b) retransmission to make certain reliable facts switch, and (c) congestion manage to minimize packet loss. Note that data facilities, therefore, best put in force quit to end drift control and there is no mechanism to ensure point to point (i.e. Link level) packet shipping as Ethernet handiest affords quality attempt service.

Over-subscription: Many data center designs introduce over-subscription as a way to lower the total fee of the design. We define the term over-subscription to be the ratio of the worst-case practicable mixture bandwidth some of the give up hosts to the total bisection bandwidth of a selected communication topology. An over-subscription of 1:1 shows that each one hosts may also doubtlessly communicate with arbitrary different hosts at the entire bandwidth in their network interface. An oversubscription cost of 5:1 method that simplest 20% of available host bandwidth is available for a few communication styles.

III. THE PROPOSED APPROACHES

Fat-tree: Al-Fares et al. [1] introduces Fat-tree, as seen in Figure 2, that permits the usage of inexpensive commodity community elements for the architecture. All switching elements within the community are indentical. Also, there are always a few paths to the cease hosts so as to use the total bandwidth. Further, the price of Fat-tree community is much less than conventional.

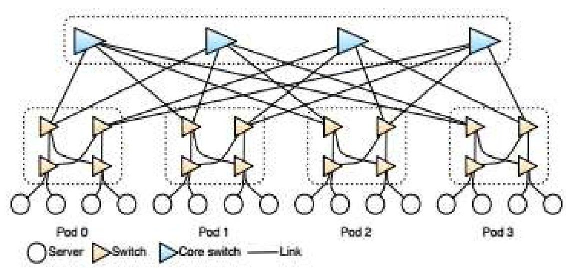

Figure 2: Fat-tree design [1].

By [1], the character of IP/Ethernet is to set up connection among source and destination spot the usage of single routing direction. Single routing path ends in predominant performance problems in Fat-tree layout. To prevent the performance troubles this design proposes two-level routing tables, which may be applied in hardware using Content-Addressable Memory (CAM). According to Al-Fares et al. [1], the size of Fat-tree relies upon on the switch residences. Switch with 48 ports can support a network with 27,648 hosts and scaling out to support networks with over a hundred,000 hosts calls for advanced switches. In addition, wiring can be very severe venture with Fat-tree design, however packaging and location strategies are proposed for this difficulty.

Monsoon: In [9] Greenberg et al. Proposes a blueprint known as Monsoon, a mesh-like structure for "cloud"-offerings that uses commodity switches to reduce the cost and lets in effective scaling over to 100,000 servers. Monsoon improves overall performance by means of using Valiant Load Balancing (VLB). Figure. 3 illustrates an outline of the Monsoon structure. The architecure is split in to Ethernet layer 2 and IP layer 3, but Monsoon makes a speciality of the layer 2. The advantages of layer 2 consist of cost financial savings, elimination of the server fragmentation (all packages can percentage a big flat cope with area) and fending off disturbance of the IP-layer capability.

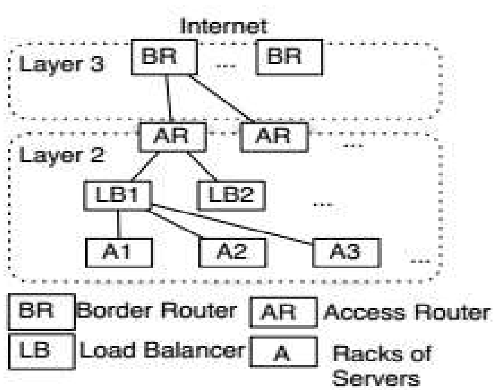

Figure 3: Monsoon design [9].

By [9], Monsoon requires layer 2 switches to have programmable control plane software, MAC-in-MAC tunneling and 16K MAC entries. Also, top-of-rack switch should handle 20 server's 1-Gbps link onto 2 10-Gbps uplinks. The upper layer switches should have 144 ports with 10-Gbps. This architecture allows over 100,000 servers with no oversubscribed links in layer 2. The load balancers (LB) can be built from commodity servers, instead of specialized and expensive hardware. IP layer 3 is responsible for dividing requests from Internet equally to access routers (AR) by Equal Cost MultiPath (ECMP).

According to Greenberg et al. [9], networking stack of a server requires replacing ARP with a user-mode process called Monsoon Agent and encapsulator, which is a new virtual Mac interface that encapsulate Ethernet frames. The Monsoon networking stack needs path information from a Directory Service. There are several ways to implement the Directory Service. Another service needed for the Monsoon design is Ingress Server, which works with Access Routers (AR). Ingress Server is required for Monsoon load spreading and encapsulation for the VLB.

BCube, MDCube: BCube [10] is a delivery-container based totally on modular statistics center (MDC) layout. MDCs are shaped by way of a frew thousands of servers which are interconnected through switches that is then packed into a 20- or 40-feet transport-box. MDC offers brief deployment time, decrease cooling and manufacturing price, and higher system and power density. Shipping field primarily based products are already presented by fundamental companies in the area, which includes HP, Microsoft and Sun.

MDCube [14] is a shape to assemble mega-data centers primarily based on boxes. Containers in MDCube comply with the BCube design, which connects thousands of servers inner the container. In other words, MDCube is a design to acquire a mega-data middle the usage of BCube-based totally packing containers as constructing blocks.

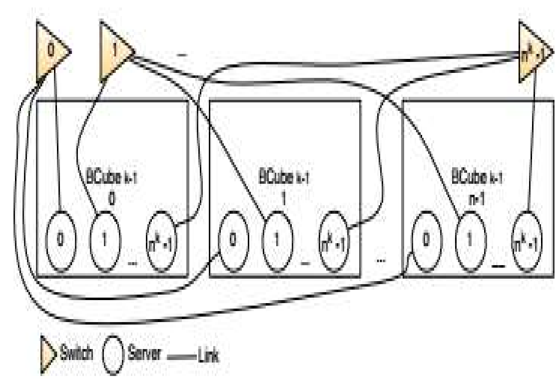

Figure 4: BCube design [10].

Figure 4 illustrates BCube [10] server centric design, which uses handiest commercial-off-the-shelf (COTS) switches and commodity servers. Each server has small range of community ports, that connect with mini-switches. The routing intelligence is left for the server. The authors claim Clos topology based totally solutions, inclusive of Monsoon, VL2 and Fattree do not guide one-to-x (one-to-one, one-to-several and one-to-all) properly, in evaluation to BCube. In addition, consequences display that BCube offers greater graceful overall performance degradation than traditional community architectures.

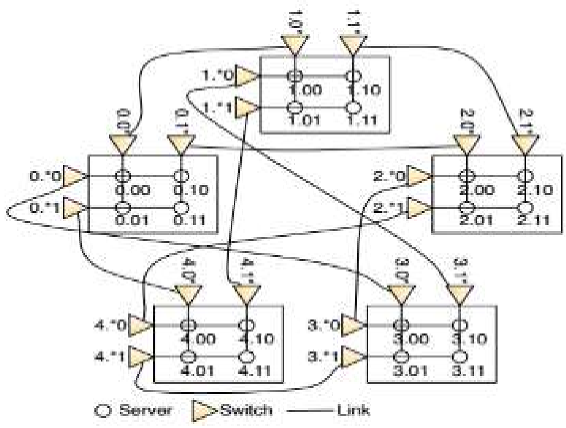

Figure 5: Example MDCube design [14].

Greenberg et al. [8] introduces VL2, a network structure that uses Valiant Load Balancing (VLB) for site visitors spreading, cope with decision helping big server pools and flat addressing to keep away from fragmentation of sources. The real topology gives route diversity. Overall VL2 is promised to solve many modern-day problems with the aid of offering agility, because it creates an phantasm of a single whole data cente layer-2 transfer by means of creating a virtual layer. Also, VL2 gets rid of the need for overscubscriping hyperlinks inside the community via the network design.

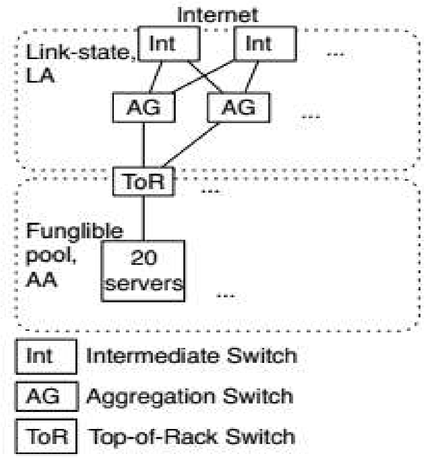

Figure 6: VL2 design [8].

Apparently in addition to Monsoon, through [8] VL2 requires a listing service and server agent for VL2 addressing and routing. Also, it seems VL2 requires adjustments to servers' network stacks to permit VL2 adddressing and routing design. Key concepts in VL2 addressing and routing are software-precise addresses (AAs) and location-precise addresses (LAs) that are used to split server call from locations, hence presenting agility. LAs are assigned for all switches and interfaces, while AAs are simplest used in packages. According to Greenberg et al [8], one VL2 design precept is to allow implementation on current hardware, in order that VL2 might be taken in use even nowadays. The authors evaluated VL2 overall performance by way of a operating prototype. The results imply that VL2 is efficient and achieves high load balancing equity. In addition, rough value estimates additionally indicate that a common community without oversubscribed links cost 14 instances greater than equivalent VL2 network.

IV. CONCLUSION

In this paper we first introduced the cutting-edge fashion inside the data center industry among some facts about the value shape. Next, we explained the troubles with todays facts middle network architectures, such as scalability, bodily constraints, useful resource oversubscription and fragmentation, reliability, utilization, fault tolerance, cost and Incast. Last, we added some lately proposed solutions for the troubles. We covered Monsoon, VL2, Fat-tree and MDCube. Each of the proposed solutions had their strengths and weaknesses, but our brief bet is that those more favorable for the industry are the ones which are deployable even these days and require minimum effort for the prevailing hardware.

REFERENCES

[1] M. Al-Fares, A. Loukissas, and A. Vahdat. A scalable, commodity data center network architecture. SIGCOMM Comput. Commun. Rev., 38(4):63–74, 2008.

[2] T. Benson, A. Anand, A. Akella, and M. Zhang. Understanding data center traffic characteristics. In WREN '09: Proceedings of the 1st ACM workshop on Research on enterprise networking, pages 65–72, New York, NY, USA, 2009. ACM.

[3] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker. Ethane: taking control of the enterprise. SIGCOMM Comput. Commun. Rev., 37(4):1–12, 2007.

[4] Y. Chen, R. Griffith, J. Liu, R. H. Katz, and A. D. Joseph. Understanding tcp incast throughput collapse in datacenter networks. In WREN '09: Proceedings of the 1st ACM workshop on Research on enterprise networking, pages 73–82, New York, NY, USA, 2009. ACM.

[5] Cisco Systems. Data center: Load balancing data center services. 2004.

[6] P. Costa, T. Zahn, A. Rowstron, G. O'Shea, and S. Schubert. Why should we integrate services, servers, and networking in a data center? In WREN '09: Proceedings of the 1st ACM workshop on Research on enterprise networking, pages 111–118, New York, NY, USA, 2009. ACM.

[7] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. The cost of a cloud: research problems in data center networks. SIGCOMM Comput. Commun. Rev., 39(1):68–73, 2009.

[8] A. Greenberg, J. R. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta. Vl2: a scalable and flexible data center network. In SIGCOMM '09: Proceedings of the ACM SIGCOMM 2009 conference on Data communication, pages 51–62, New York, NY, USA, 2009. ACM.

[9] A. Greenberg, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta. Towards a next generation data center architecture: scalability and commoditization. In PRESTO '08: Proceedings of the ACM workshop on Programmable routers for extensible services of tomorrow, pages 57–62, New York, NY, USA, 2008. ACM.

[10] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang, and S. Lu. Bcube: a high performance, server-centric network architecture for modular data centers. In SIGCOMM '09: Proceedings of the ACM SIGCOMM 2009 conference on Data communication, pages 63–74, New York, NY, USA, 2009. ACM.

[11] M. Bari, R. Boutaba, R. Esteves, L. Granville, M. Podlesny, M. Rabbani, Q. Zhang, and M. Zhani, "Data center network virtualization: A survey," Communications Surveys Tutorials, IEEE, vol. PP, no. 99, pp. 1 –20, 2012.

[12] C. Kachris and I. Tomkos, "A survey on optical interconnects for data centers," Communications Surveys Tutorials, IEEE, vol. 14, no. 4, pp. 1021 –1036, quarter 2012.

[13] K. Wu, J. Xiao, and L. Ni, "Rethinking the architecture design of data center networks," Frontiers of Computer Science, vol. 6, pp. 596–603, 2012. [Online]. Available: http://dx.doi.org/10.1007/s11704-012-1155-6

[14] H. Zimmermann, "OSI reference model–The ISO model of architecture for open systems interconnection," Communications, IEEE Transactions on, vol. 28, no. 4, pp. 425–432, 1980.

[15] K. Kant, "Data center evolution: A tutorial on state of the art, issues, and challenges," Computer Networks, vol. 53, no. 17, pp. 2939–2965, 2009.

BIOGRAPHY

Dr.K.Praveen Kumar received the PhD in Computer Science & Engineering in 2015, M.Tech in Software Engineering from Kakatiya Institute of Technology & Science Warangal, Telangana, India in 2010 and B.Tech in Information Technology from Kakatiya

Institute of Technology & Science Warangal, Telangana, India 2007. Presently working as Assistant Professor in Computer Science Department at Adama Science and Technology University, Adama, Ethiopia.

K Sree Ranganayaki pursuing M.Tech in Computer Science & Engineering at Chaitanya Institute of Technology and Science. Her Area of Interest is Network Science and Cloud Computing.