ID:1407870 Name: Muhan Guan Username: Muhan GUAN

Survey on Networking in Cloud Data Centres

1. Introduction

Large-scale data centers always act as a bedrock to provide support for a considerable number of cloud computing services (i.e., remote meeting and social networking) and infrastructure services (i.e., GFS, MapReduce and HDFS). However, the network connecting millions of servers in data centers plays a critical role in the whole system, especially when those various popular cloud-based services are generated, which presents much higher requirements to capability of network than ever, such as reliability and scalability. Therefore, the research into various aspects of networks in cloud data centers has long been an inevitable challenge and accordingly a crucial topic.

In conventional distributed computing architectures, specific applications are exclusively served by physically dividing network resources. In contrast, cloud service data center networks (DCNs) apply virtualization technologies on the allocation of various applications across multiple Virtual Machines (VMs) [1]. This process involves an advanced degree of autonomous resource management, facilitated by a range of virtualization software. This network responds effectively to variations in demand and alterations in network conditions, thereby optimizing resource usage and enhancing the quality of services provided.

To enhance user satisfaction across various cloud services, enterprises are predominantly confronted with two pivotal challenges. Firstly, the development of scalable and cost-effective architectures of DCNs is essential, which can support large server populations while maintaining efficient bidirectional bandwidth capacities among network components. Secondly, through the astute application of virtualization technologies and routing algorithms, it's crucial to optimize DCNs' resource utilization and augment the volume of services delivered using identical resource quantities, all while upholding a resilient network state that can adequately withstand link or server failures. [2,3]

The structure of rest of this survey is as follows: Section 2 presents an overview of three main research areas about networking in Cloud Data Centers, while Section 3 compares three key physical DCNs structures. Finally, Section 4 provides a conclusion and potential future research topics about networking in Cloud Data Centers.

2． Related work

As mentioned above, research related networking in could data centers can be broadly segmented into four paramount facets: the physical architecture of the DCN, the techniques of virtualization, and the routing protocols and traffic engineering.

2.1 Physical Architectures

The architecture of a DCN is deemed as an essential determinant because of its pivotal role in ensuring the reliability of a Data Center. [4] In light of the drawbacks and limitations of traditional Three-Tier network architectures, including increased network costs and significant bandwidth oversubscription towards the core tier, considerable research has been undertaken to identify an efficient interconnection architecture for cloud data center networks. This has resulted in three predominant categories: Server-Centric, Switch-Centric, and the recently proposed Dual-centric architectures.

A. Switch-Centric Architectures

Within the switch-centric architecture, network intelligence is predominantly situated on the switches. Each server typically engages a single Network Interface Card (NIC) port to connect to a Top-of-Rack (ToR) switch, which interconnects via aggregation switches linked through core switches. In these schemas, servers function as endpoint hosts for data transmission and reception, with switches managing end-to-end routing, two classic switch-centric schemes are Fat-Tree [5] and V L2[6].

B. Server-Centric Architectures

The server-centric architecture, a category of DCNs, positions network intelligence primarily on servers, which function as both forwarding and computing units. In this structure, exemplified by DCell[7] and BCube [8] servers employ multiple NIC ports to connect to various switch layers and potentially other servers. Switches, if present, typically serve as basic crossbars. Servers assume dual roles in this architecture, functioning as end hosts while also acting as relays for other servers, with their core task being the selection of the next forwarder in the network.

C. Dual-centric architectures

In the dual-centric architecture, the task of routing intelligence is distributed between servers and switches. This configuration can concurrently harness the high programmability of servers and the significant switching capability of switches. This architecture facilitates all permutations of connectivity, encompassing server-to-server, server-to-switch, and switch-to-switch connections. Key examples are RibsNet[4] and FCell[9], which are 2 recent advancements.

Design principles for physical architectures

Primarily, the paramount objectives for efficacious architectures of DCNs that can support tremendous cloud services encompass following points. [9]

Fault-tolerance: An optimally designed data center network topology, bolstered by the principle of fault-tolerance, prioritizes the integration of redundant elements and connections, thereby creating multiple concurrent linkages among nodes, mitigating the risk of single-point disruptions, and ensuring the preservation of service quality even in the face of unexpected network perturbations.

Scalability: The advent of numerous emergent cloud services necessitates that DCNs are designed to accommodate a burgeoning number of servers within their infrastructures. The integration of novel elements into the network should not impede the efficient functionality of extant servers within the system.

Bandwidth: Given the substantial data transmission requirements, data centers necessitate a network architecture with considerable bandwidth to accommodate high-throughput transfers, and the capability to automatically manage unexpected surges in workload. Acknowledging that server-to-server traffic often surpasses that between the data center and external clients, it becomes essential to ensure abundant bandwidth within every server pair in the DCN structure.

Cost: The costs of cloud data center networks originate from various aspects, encompassing hardware, supplementary infrastructure, and computing resources. Consequently, when developing extensive data center networks, the consideration of cost investment becomes especially crucial.

2.2 Routing schemes and traffic engineering

Routing schemes

Routing protocols play a crucial role in the efficiency of Cloud Data Center Networks. Each network architecture exhibits distinct characteristics, and encompassing its individual topology, hence they have their own specific routing algorithm and traffic control solution. Unlike traditional Internet link-state routing protocols, the majority of prevailing data center network routing protocols are tailored explicitly for specific network topologies. Broadly, the interconnection structures of DCNs can be primarily divided into two principal schemes, which are routing in server-centric structure and routing in switch-centric structure respectively.[12]

Traffic engineering

Although rudimentary routing schemes primarily pursue low-latency paths between servers, advanced DCN routing necessitates comprehensive optimization encompassing latency, reliability, throughput, and energy efficiency, among other factors. This multidimensional optimization represents the traffic engineering (TE) problem.

The traffic in DCNs is divided into two primary parts, inter-DCN and intra-DCN traffic. Given the predominance of within-data-center communications in overall DCN activity, the overall performance of a DCN largely rely on inter-DCN TE.

It's noteworthy that numerous scholars have employed various sophisticated machine learning-based models for traffic forecasting, enabling precise real-time prediction and scientific control of intra-cloud data center traffic. Li et al. [13] conducted research on traffic transmission within data centers, integrating the wavelet transform technique with a neural network. Additionally, they utilized the interpolation filling method to mitigate the monitoring overhead instigated by the irregular spatial distribution of data center traffic.

Design principles for TE

In the process of designing a TE model, the following design principles should be considered [1,2].

Reliability: The first objective both providers and subscribers in DCNs, is paramount in ensuring robust data transmission for essential services and business operations. Effective TE models optimize routing patterns, enhancing network resilience through multi-pathing and fault-tolerance, thereby rendering the DCNs reliable, robust, and indispensable to its users.

Load-balancing: The second objective is to optimize link capacity usage to address the latency-throughput trade-off. With a multitude of cloud applications and services running concurrently in a data center, an intelligent routing protocol should ensure optimal performance for each by judiciously distributing traffic across intra-data center links, thereby maximizing capacity utilization and achieving even traffic dispersion.

Power expenses: A critical objective for competitive service pricing in DCNs. Thus, proficient TE models should drive routing strategies to minimize the usage of links and switches, reducing energy consumption. This approach not only maximizes profits for DCNs but also enables them to maintain market-competitive service costs.

2.3 Virtualization techniques

Virtualization techniques, mainly including Server Virtualization (SV), Network I/O Virtualization (NIOV), Network Virtualization (NV), and Resource Virtualization (RV), enable the formation of a virtualized infrastructure within a DCN. This infrastructure, composed of both wholly and partially virtualized elements like servers, links, switches, and topology, partitions and allocates resources to users, enhancing the suitability of DCNs for cloud-based services.[16]

Furthermore, employing virtualization techniques, DCNs can deliver enhanced reliability and performance for data and applications, while simultaneously reducing costs, irrespective of their geographical location.

SV: Utilizing software or firmware known as a hypervisor, physical hardware is partitioned into multiple insulated and autonomous virtual instances, thereby addressing the challenge of Virtual Machine (VM) allocation in cloud computing. As a VM instance is instantiated and positioned within its host via cloud system's intelligent solutions (for instance, Swift), a DCN must streamline the interaction between disparate hypervisors to expedite the allocation process.

NIOV: NIOV offers an abstraction that allows sharing of network I/O devices among VMs. One prevalent NIOV form in cloud DCNs is NIC virtualization, where a physical NIC is shared between multiple VMs via a hypervisor. Catering to cloud computing requirements, a DCN deploys NIOV to deliver extensive bandwidth capacity and minimal CPU overhead access to each VM within the virtualized I/O device.

NV: facilitates the formulation of logically distinct virtual networks atop the physical DCN infrastructure. Hence, NV addresses the necessities of cloud computing concerning the architecture of VI topology, addressing, and routing.

RV: RV effectively tackles the demands of virtual infrastructures in terms of resource allocation and VM management. The DCN must strategize VM placement to prevent oversubscription and ensure adherence to the resources outlined in the Service Level Agreement (SLA).

3. Comparison of three Cloud Data Center Network Architecture

In this section, I will embark on a comparative analysis of three distinct cloud data center network architectures, namely the two classic structures - Fat-tree [5] and BCube [7], alongside RibsNet [4], a novel two-layer based dual-centric structure recently brought to light by AL-Makhlafi et al.

3.1 Fault Tolerance: The Fat-tree network architecture excels in providing multiple redundant transmission paths at all levels, devoid of any singular point of failure, thus bolstering superior fault tolerance capabilities. The interconnection methodology and recursive structure inherent to BCube facilitate modular construction of the data center, leading to effective implementation. Particularly in scenarios involving incomplete structures, BCube surpasses Fat-tree, demonstrating stable and robust fault-tolerant performance. RibsNet, on the other hand, introduces an exclusive fault-tolerant routing protocol, referred to as RFT [4], which adeptly addresses failures associated with switches and links within the network.

3.2 Scalability: Furthermore, the scalability potential of the Fat-tree network is somewhat inhibited due to the inherently limited number of physical ports available at its switches. BCube, by virtue of its recursive methodology, exhibits an impressive capacity for supporting large-scale servers, thereby providing convenience for expansion. This inherent trait facilitates modular construction and confers notable network scalability. Contrasted with Fat-tree and BCube, RibsNet presents a compelling argument for large-scale DCN construction, given its capability to accommodate an increased volume of servers while maintaining network symmetry. Moreover, its highly incremental scalability mitigates the issue of port idleness, rendering RibsNet a particularly optimal solution for large-scale DCN endeavors.

3.3 Bandwidth: As the network scale expands, the Fat-tree structure showcases superior throughput. It efficiently achieves load balancing among multiple links at the core level, thereby circumventing potential network performance bottlenecks. BCube network, noted for its substantial bandwidth, capably fulfills the requirements of high-throughput data transmission. Empirical evidence underpins the assertion that, under equivalent construction costs, RibsNet provides a higher available network bandwidth.

3.4 Cost: The Fat-tree topology leverages identical switches, thus enabling the utilization of affordable commodity switches and subsequently facilitating cost-effective DCNs. Conversely, due to BCube's structural design that incorporates direct connections between network switches and servers, it invariably incurs substantial wiring expenses. RibsNet exhibits a significant cost advantage over both Fat-tree and BCube. Owing to its unique structural attributes, the construction of large-scale data centers with RibsNet requires fewer switches and wires, considerably reducing not only its cost but also the complexity of the overall network structure.

4. Conclusions and Future Directions

4.1 Conclusion

This survey initially presents the pertinent work in three critical research directions within the domain of cloud data center networks, introducing their respective crucial characteristics. Subsequently, I compare two classic network structures with a newly proposed one, finding that RibsNet outperforms the other two structures across four key evaluation metrics. Lastly, the paper identifies two potential avenues for future research.

4.2 Future Direction

A. ML-based techniques on TE

At present, machine learning technologies such as deep neural networks and reinforcement learning algorithms have been integrated into various domains within cloud data center networks. However, the progress across these domains is notably disparate. Predominant research is primarily concentrated on flow prediction, resource management, and flow classification [13]. Nonetheless, there appears to be a less pronounced focus on route optimization and congestion control, areas that are integral for enhancing the overall intelligence of the network.

B. DNC Migration

Additionally, to cater to load and scale requirements, a Virtual DCN (VDCN) provider might opt to transition from its current location (that is, the physical DCN currently hosting the VDCN) to a new location (a new physical DCN) offering larger resource or scale capacities to meet the dynamic VDCN load demands. Research in this realm can be considered a novel and burgeoning direction that is still in its nascent stages.

5.Reference

[1] Ahmad Nahar Quttoum (2018) 'Interconnection Structures, Management and Routing Challenges in Cloud-Service Data Center Networks: A Survey', International Journal of Interactive Mobile Technologies, 12(1), pp. 36–60. doi:10.3991/ijim.v12i1.7573.

[2] Wang, B. et al. (2015) 'A survey on data center networking for cloud computing', Computer Networks The International Journal of Computer and Telecommunications Networking, 91, p. 528. doi:10.1016/j.comnet.2015.08.040.

[3] Li, Y. et al. (2023) 'A Weighted Optimal Scheduling Scheme for Congestion Control in Cloud Data Center Networks', IEEE Transactions on Services Computing, pp. 1–9. doi:10.1109/TSC.2023.3239524.

[4] AL-Makhlafi, M. et al. (2022) 'RibsNet: A Scalable, High-Performance, and Cost-Effective Two-Layer Based Cloud Data Center Network Architecture', IEEE Transactions on Network and Service Management, Network and Service Management, IEEE Transactions on, IEEE Trans. Netw. Serv. Manage, PP(99), p. 1. doi:10.1109/TNSM.2022.3218127.

[5] Al-Fares, M., Loukissas, A. and Vahdat, A. (2008) 'A scalable, commodity data center network architecture', ACM SIGCOMM COMPUTER COMMUNICATION REVIEW, 38(4), pp. 63–74. Available at: https://discovery.ebsco.com/linkprocessor/plink?id=1d212bc9-c783-3a5c-9b0d-40ed7b4df649 (Accessed: 17 May 2023).

[6] Greenberg, A. et al. (2011) 'VL2: A Scalable and Flexible Data Center Network', Communications of the ACM, 54(3), pp. 95–104. doi:10.1145/1897852.1897877.

[7] Guo, C. et al. (2008) ‘Dcell: A scalable and fault-tolerant network structure for data centers’, Computer Communication Review, 38(4), pp. 75–86. doi:10.1145/1402946.1402968.
[8] Guo, C. et al. (2009) ‘BCube: A High Performance, Server-centric Network Architecture for Modular Data Centers’, ACM SIGCOMM COMPUTER COMMUNICATION REVIEW, 39(4), pp. 63–74. Available at: https://discovery.ebsco.com/linkprocessor/plink?id=a60902b7-35ce-3e5a-92cd-12f9dcd39e0d (Accessed: 18 May 2023).
[9] Li, D. et al. (2017) ‘Towards the Tradeoffs in Designing Data Center Network Architectures’, IEEE Transactions on Parallel and Distributed Systems, Parallel and Distributed Systems, IEEE Transactions on, IEEE Trans. Parallel Distrib. Syst, 28(1), pp. 260–273. doi:10.1109/TPDS.2016.2610970.
[10] Xie, J. et al. (2017) ‘An Incrementally Scalable and Cost-Efficient Interconnection Structure for Data Centers’, IEEE Transactions on Parallel and Distributed Systems, Parallel and Distributed Systems, IEEE Transactions on, IEEE Trans. Parallel Distrib. Syst, 28(6), pp. 1578–1592. doi:10.1109/TPDS.2016.2629508.
[11] Talebian, H. et al. (2020) ‘Optimizing virtual machine placement in IaaS data centers: taxonomy, review and open issues’, Cluster Computing: The Journal of Networks, Software Tools and Applications, 23(2), pp. 837–878. doi:10.1007/s10586-019-02954-w.
[12] Kai Chen et al. (2011) ‘Survey on routing in data centers: insights and future directions’, IEEE Network, Network, IEEE, 25(4). doi:10.1109/MNET.2011.5958002.
[13] Li, B. et al. (2022) ‘Machine Learning Empowered Intelligent Data Center Networking: A Survey’. Available at: https://discovery.ebsco.com/linkprocessor/plink?id=fb9bd338-a984-3925-bd8c-c63014c3fdd6 (Accessed: 17 May 2023).