Legacy Data Center Network (DCN) Reviewed: considering Big Data Stream Mobile Computing (BDSMC) applications

¹Godspower I. Akawuku, ²Rapheal O. Okonkwo, ³Keneddy C. Okofor ^{1,2}Nnamdi Azikiwe Unversity, Awka (NAU), Nigeria. ³Federal University of Technology, Owerri (FUTO), Nigeria. **ABSTRACT**

Recently, Enterprises that operate over vast geographical areas uses multiple data centers to collect, store and process data in real time via energy-efficient acquisition, wirelessly transport clients or users data to the cloud. This paper seeks to technically review existing Data Center Networks (DCN) considering BDSMC applications. In this research, related works on Distributed Data Center Networks will be presented. Within the stream computing ecosystems, there are various network models but the pool of possible DCN topologies/architectures to adopt appears little and unfit for the purpose of BDSMC optimization. However, most related Data center architecture will be reviewed. Investigate efforts on both server centric and switch-centric models adaptable to BDSMC network layer. The extent of work done on distributed spine-leaf re-designed server-centric network construction so as to automatically harvest network interconnection into a 'stellar' dual-port server-centric SG network; how classical graph-based interconnection network translate network performance similar to generic works for BDSMC ecosystems. Review stellar transformation using the well-studied generalized hypercube family of interconnection networks for BDSMC ecosystems. The literature was searched from the databases: IEEE Xplore Digital Library, Springer Link Digital Library, and Google Scholar, IET Digital Library, Frontiers Library, ACM Digital Library repositories resulting in 98 papers after several eliminations ranging from year 2000-2022. In conclusion, state-of-the-art dual-port server-centric DCNs (FiConn, DCell, DPillar), etc, while looking at possible architectures with excellent comparative performance for BDSMC ecosystems. Research gaps are revealed for further study.

Keywords: Data, Center, Network, Data Stream, Mobile Computing and applications.

INTRODUCTION

Large scale organizations that operate over vast geographical areas uses multiple data centers to collect and store data from their clients or users by [1] For example, most transport companies collect data from different departments that operate in remote locations. The streamed data is usually stored in multiple data centers, and its volume grows rapidly, often exceeding terabytes in each data center in the works of [2, 3]. In many cases, these organizations analyzes their big data in multiple data centers in a batch for analytics and business decisions widely discussed in [4]. When taking data in multiple data centers into consideration as

a whole, analyzing such big data efficiently and effectively becomes a big challenge for service offloading. Apache [5, 6, 7, 8, 9] are the most popular frameworks for parallel and distributed big data stream processing and analysis says [10]. However, these frameworks are designed to process data locally within the same data center. Hence, they are in need to copy all data to a single data center before processing a locally distributed computation in [11].

BDSMC explains a new generation of mobile or wireless integrated computational networking infrastructures designed to extract hidden value from an ever-increasing volume of space-time

correlated heterogeneous data streams in [12]. This is enabled in real time via energy-efficient acquisition. wireless transport, and processing. The characterization of the BDSMC paradigm expresses (i.e., varietv (data heterogeneity), volume (ever increasing amount of data to be processed), velocity (data generation at fast and unpredictable rates), value (huge value but hidden in massive datasets at very low density), and volatility (the acquired data streams must

Figure 2.0 illustrates a typical Fat-tree architecture widely used in most enterprise networks. In [13] the authors how to leverage presented largely commodity Ethernet switches to support the full aggregate bandwidth of clusters consisting of tens of thousands of elements. A simple topology was studied for streams propagation in Figure 1.0 which is not useful in BDSMC ecosystems depending on the scale. Most legacy

Akawuku et al

be transported and processed in real time). While the first four Vs are common to all big data applications, another last V (i.e., volatility) is introduced for featuring big data stream applications. In general, the value of a stream of data is closely related to both its space and time coordinates, and hence, after acquisition, this value quickly decreases if the computing-pluscommunication delay is larger than a suitable quality of service (QoS)-dictated hard threshold.

Non-BDSMC Fat-Tree Architecture

networks use Fat-tree data structure in which top branches have stronger/thicker branch than others in the design hierarchy. These data links branches vary in thickness (bandwidth) while allowing the links for more efficient and technologyspecific use. Mesh and hypercube topologies uses common design requirements based on rigid algorithm.



Figure 1.0: Fat-tree Architecture [13].

[14] a special case of Clos network, also referred to as a folded Clos network was depicted as fat-tree data center topology. Originally, the fat-tree topology principle was suggested for supercomputing and was modified for data center networks. In terms of the number of pods, a fat-tree topology is usually referred. This is numerated from Pod-0 to Pod-(k-1) from left to right. The topology consists of *k* pods with three switch layers: edge switches, switches for aggregation, and switches for the center. Thus, there are k switches (each with k ports) in each pod arranged in two layers of k/2 switches in kpod fat-tree topology, one layer for edge switches and the other for aggregation switches. The k/2 aggregation switches are attached to each edge switch. (k/2)2 core switches are available, each of which connects to k pods. The topological knowledge description is shown in Table 1.0.

www.idosr.org

Number of Pods	К
Number of middle (Core) Switches	(k/2) ²
Number of gateway Switches	K ² /2
Number of Edge end Switches	K ² /2
Number of rack Switches, N (all types)	5K ² /4
Number of connection links, L	K ³ /2
Number of Supported Hosts	K ³ /2

Table 1.0: Fat-tree Topology

For k = 4, Figure 2 shows the fat-tree topology. Here, the bottom 8 edge switches (nodes), numbered 1 to 8, have 16 $MPM = Min\left\{\frac{L-1}{k}, 1.0\right\}$

supported physical hosts. This is known as the outcome of
$$K + L$$
. Leading to multipath measure (MPM) in Equ 1 [15].

1.



Figure 2: DCN topology of a 4-pod fat-tree [15].



Figure 3: A 6-pod fat-tree DCN topology [15].

There are some advantages of fat-tree topologies. With the same number of ports, all switches are of the same form, with each port usually having the same speed;

the end hosts often support the same speed. Between any two hosts, there are several routes. It is important to link each host to an edge switch first. There are two paths between two edge switches within a

pod (intra-pod) in a 4-pod fat-tree topology and there are four paths between any two edge switches that are across pods. This model is not scalable for BDSMC ecosystem. Various efforts in literature seem to isolate BDSMC application. For instance, the work of [16] focused on

Non-BDSMC Dahu Datacenter Architecture

Dahu DCN was described in [21] as an improvement on networks powered by commodity Ethernet switches which direct supports link networks. Bv dynamically distributing traffic equally across links, this model eradicates congestion points. When performing load balancing using local data, it does traffic forwarding over-non-minimal routes. HyperX topology, a direct connection network for detailing how Dahu's hardware

Akawuku et al

active job placement and routing algorithms in the Fat-Tree Topology. Similarly, the works [17, 18, 19, 20] explored the various deployments of fattree without any application in BDSMC system.

primitives are used and tested has been studied. A related work on HyperX topology was addressed in [22], the architecture of which is shown in Figure 3. HyperX is an *L*-dimensional direct network between any server pair with several paths of varying length. It can be interpreted as a Hyper-Cube topology generalization. The Hypergraph architecture is scalable but not optimized for BDSMC application.



Figure 4: HyperX topology (*L=2, S=*3)

Akawuku et al



Figure 5: Dahu Architectural Data-path pipeline for packet forwarding [23]. Dahu new switch DCN mechanism for dynamic traffic hashing allows various network paths to be hashed and exploits non-shortest path forwarding to minimize congestion while using novel hardware switch primitives and control software to forwarding avoid persistent loops. However, the decentralized load balancing heuristic is not scalable for BDSMC for environments (i.e., rapid local decisions to alleviate congestion. BDSMC

center implementations data would obviously be primarily restricted to multirooted tree switch-centered topologies such as [24] as illustrated in Figure 5. Intelligent switches are used in Switch-Centric category to conduct smart packet routing in a Data centers. In this category, some switch-centric data center topologies include: Clos-Network in the works of [25, 26, 27, 28, 29, 30], among others. These will be further reviewed in this Section.



Figure 6: VL2 Architecture [31].

Non-BDSMC Clos Network

Akawuku et al

The authors [32], addressed the design of a three-stage buffered Clos-network switch (TSBCS) along with a new mechanism for batch scheduling (BS). The TSBCS/BS is

mapped to a "fat" combined input-crosspoint queuing (CICQ) switch with a directly implemented CICQ scheduling algorithm in TSBCSS.



Figure 7: A Clos switch architecture of three-stage shared-memory.Similarly, in [33], a parallel WavelengthTolerant Clos-network (FTC) planes, asFault Tolerant Clos network, PW-FTC, is
addressed. This was used to achieve Faultshown in Figure 8, a wavelength switch is
performed.



Figure 8: Clos Network with Parallel wavelength fault tolerance [34],

Obviously, such traditional buffered Clos networks are not attractive to large-scale packet switches needed in BDSMC ecosystems due to the out-of order packets or local congestion and underutilization problems.

Jelly Fish Architecture

In the top-of-rack (ToR) switch layer, the Jellyfish DCN schemes are based on the construction of a random graph. In this case, there are a number of ports for each ToR switch that connect to other ToR switches, while the remaining server ports are used by [35].

Akawuku et al



Figure 9: Jellyfish Architecture [36]. Datacenter Optical Switch (DOS) Architecture

In this section. another class of sophisticated DC models will be reviewed. The authors [37], presented the Datacenter Optical Switch (DOS) designed for scalable and high-throughput data center interconnections, but not for the BDSMC device. The architecture is based on an arrayed waveguide grating router (AWGR) that enables the wavelength domain to

overcome contention. The likelihood of lower latency and higher throughput even at high input loads has been illustrated by its switching architecture. DOS has however, not been used in any known BDSMC systems. To date, this has never been extended to BDSMC data center networks.





(OLG: Optical Label Generator; PE: Packet Encapsulation; LE: Label Extractor; FDL: Fiber Delay Line; PFC: Packet Format Converter; O/E: Optical-to-Electrical Hybrid Packet/Circuit Converter; E/O: Electric-to-Optical Converter; TX: Transmitter; RX: Receiver (i): Node I Label

Hybrid Packet/Circuit Datacenter Architecture

Hybrid Packet/Circuit (HYPAC) switched DCNs were introduced optical Fibre communication DCN owing to its merits of bandwidth capacity and power efficiency says [39]. Figure 2.10 uses their proposed Collaborative Bandwidth Allocation (CBA), which optimally sets the network up. A rack-to-rack optical circuit switched (OCS) network complements the electric packet switched (EPS) network.



Figure 11: Hypac DCN Architecture [40]. Flattened Butterfly Architecture

The authors [41], presented a flattened butterfly for high-radix networks as a costeffective topology in Figure 12. This is useful in load-balanced traffic where, relative to the Clos network, its efficiency is very optimum. The advantage over the Clos is achieved by removing redundant hops when they are not required for load balancing. It uses high radix flattened butterfly topology that offers greater route



diversity than a traditional butterfly and has around half the cost on balanced traffic of a comparable Clos efficiency network. This network is used when the level or radius of interconnection networks is increased by an increase in integrated-circuit pin bandwidth. Again, the network integration of BDSMC systems may encounter computational overhead here.



Figure 12: Flattened Butterfly Architecture [42].

(a) 4-ary 2-fly butterfly and (b) 4-ary 2-flat - the corresponding flattened butterfly with a single dimension, (c) 2-ary 4-fly butterfly and (d) 2-ary 4-flat - the corresponding flattened butterfly with FRINGE Arc

The research in [11], used Software-Defined Networking (SDN) as a smart tool, called FRINGE, to derive an effective software-define edge control system. The concept is to boost the Ethernet DCN's scalability. SDN domains on commodity access switches at the edge of the DCN are three dimensions. Lines denote unidirectional links in the butterfly and bidirectional links (i.e. two unidirectional links) in the flattened butterfly.

FRINGE Architecture

derived from the FRINGE architecture i.e.: Top-of-Rack (ToR) switches. The value is that it can be extended to DCN topologies that are arbitrary. It has access switches, recognized OpenFlow-enabled ToR switches (OFToR) at rest, or installed at the edge of DCN. The other portion of FRINGE

Akawuku et al

is the upper DCN for legacy switches/routers. The logical centralized SDN controller is used in the design to

manage DCN states, such as mapping connections between OFToRs and hosts.



Figure 13: FRINGE framework [43]. Layered Scalable Data Center Architecture

The authors introduced a data center topology in [44] called LaScaDa (Layered Scalable Data center) for the construction of scalable and cost-effective networking infrastructures for data centers. Their architecture configures nodes in uniform clusters and then interconnects the clusters in an ordered manner with a system of coordinates for nodes just to reduce the number of redundant links between clusters. thus optimizing communication. The LaScaDa forwarding packets between nodes are shown in Figure 14 using a new hierarchical row-based routing algorithm. A LaScaDa network constructed from n-port switches is a layered and recursive topology that uses n-port switches to create a *k*-layer LaScaDa network (k > 1) by interconnecting n³² (k-1)-Layer LaScaDa networks. These switches are designated as internal switches asserted [45].

The architecture uses its algorithm to construct the route of the source on the bases on the differences between the coordinates for the source and destination. In addition, a large number of nodes are connected using a small node degree by the proposed topology.



Figure 14: LaScaDa network Architecture [46].

Akawuku et al



Figure 15: 2-layer LaScaDa architecture

The authors [48], considered low-delay switches with end-to-end latency in largescale High-Performance Computing (HPC) interconnect with cable delays. Skywalk was the new architecture created to satisfy the deployment (HPC systems). Skywalk uses randomness to optimize low latency

cture Figure 16: 2-layer LaScaDa node DCN [47] Skywalk Architecture

> which is very necessary in BDSMC network designs. This is done in a manner that accounts for the physical layout of the topology in order to cascade further cable length with latency reductions as shown in Figure 17





Clearly, vast numbers of servers must be interconnected by the BDSMC network architecture. in meeting But the requirements of BDSMC Data Center conventional tree-based Networks, architectures seem unstable and inefficient (DCNs). These architectures fall into two groups, namely: switch-centric designs and server-centric designs, taking into account whether the interconnection intelligence is provided on the switches or

on the servers (on-going). So far, it is evident that the switch functionality must expanded to meet the need for be interconnection in the switch-centric designs highlighted above, although servers do not need to be modified for interconnection purposes. As such, it will need high-end smart switches in a reliable BDSMC network. though this will significantly increase the interconnection cost.

Server-centric Datacenter Architectures

Interestingly, BDSMC systems will rely on server-centric designs to ensure robust server data stream integration. The main server centric models in literature include [50]: DCell in Chuanxiong [51, 52, 53, 54]. General Hypercube in [55], HSDC [56], and Stellar Transformation in [56]. These has been studied and discussed in this Section as part of related works.

DCell Architecture

A clear challenge in DC networking boarders on efficient interconnection of the number of servers is exponentially rising especially in the case of BDSMC. The [3] discussed the DCell as a authors recursive structure with a high-level DCell that is being built in a completely interconnected fashion from several lowlevel DCells linking other DCells at the same level. As the node degree increases, their work scales doubly exponentially. The research clarified that since it has no single point of failure, DCell is fault tolerant and may be useful in BDSMC application. Except in the case of critical extreme connection or node failures, its

distributed fault-tolerantrouting protocol typically performs similar to shortest-path routing. Illustration 18 demonstrates the DCell architecture showing the network capacity for different types of services that appears better than the conventional treebased framework. To build its recursively specified architecture, DCell uses servers equipped with multiple network ports and mini-switches. A server is linked to several other servers in DCell and a mini-switch through communication links, which are bidirectional by design. A more robust algorithm is needed for BDSMC application traffic density in DCell. This is missing from DCell legacy designs.



Figure 18: DCell Architecture [15] BCube system Architecture

The authors [14,17], discussed the BCubenetwork architecture typical built and customized Figure 19 shows modular data centers for shipping-container-based data centers. The architecture has a servercentric network structure, central or at the heart. This is where the servers connect to several layers of Commercial-off-the-shelf (CoTS) commodity mini-switches with multiple network ports. The servers act both as hosts and as relay nodes for each other in their architecture. By accelerating one-to-one, one-to-many and one-to-all traffic patterns, and by providing high network capacity for all-to-all traffic, their architecture is assisted by several

bandwidth-intensive applications. The main problem with BCube, as the server and/or switch failure rate rises, is that it demonstrates incremental performance deterioration. For BDSMC application, this is inappropriate. In fact due to the sealed and operational status that makes it extremely difficult to fix or replace its components, this behavior is important for shipping container data centers. It's metrits includes seamlessintegrated with the TCP/IP protocol stack, efficient packet forwarding in both hardware and software, fault tolerance. load balancing and bandwidthintensive application supports.



Figure 19: BCube Architecture . FiConn Datacenter Architecture

The authors presented FiConn network architecture in [25], which uses both ports and low-end commodity switches to derive a highly efficient and scalable structure. The server node degree is two in its layout in the design, but the design is scalable to include hundreds of thousands of low diameter and high bisection width servers. Many server machines are interconnected by FiConn DCN with low equipment cost, high and balanced network power, and

link/server fault robustness. Their framework appears to deliver lowoverhead traffic-aware routing mechanisms based on dynamic traffic state to enhance efficient connection utilization. Both the FiConn2 recursive and the extended recursive architectures. It can be inferred that the computational overhead in BDSMC active design is a major challenge



Figure 20a: FiConn₂ Recursive Architecture (Liu *et al.*, 2018).



Figure 20b: FiConn₂expanded Recursive Architecture. HCN & BCN Datacenter Architecture

An interesting server-centric data center network was found in [15] where many servers were interconnected based on optimized cost devices. Though expandable with smaller level of regularity and symmetry, centered on the 2nd degree server model, their model leveraged highend switching systems called hierarchical irregular compound network (HCN) and bidimensional compound networks (BCN). Low overhead and robust routing schemes large server deployment for were considered in the design. In the modular DCN architectures of data centers, the HCN offers great potential for a high degree of regularity, scalability and symmetry. The BCN provides the most scalable network structure with characteristics such as low diameter, high width of bisection, large number of one-to-one traffic node-disjoint paths, and strong fault-tolerant capacity. These architectures are ideal with desirable properties,

but for smart grid structural integrations, this represents great computational overhead.





(a) HCN Architecture

(b) BCN Architecture

Figure 21: HCN & BCN Architecture **DPillar Datacenter Architecture**

The research of [17] discussed DPillar as a massively scalable datacenter interconnection architecture that leverages low-end off-the-shelf commodity PC servers and switches. The DPillar uses minimal resource based on low-cost plugand-play layer-2 nodes whose servers uses

dual-port commodity PCs. Its major attribute is that DPillar scales to limitless number of servers without needing physical server upgrades. The structural layout is shown in Figure 22. The challenge with DPillar is the computational workload concerns with the deployed setup. This is

www.idosr.org not efficient for Smart grid integration model.





Switch-nodes are represented by Independent NB: the left-most & right-most server columns Squares; server-nodes are represented by dots. are the same (server-column 0))

MCube Datacenter Architecture

The work in [14] discussed MCube model which is a server-centric network architecture primarily designed for data centers. MCube uses commodity mini switches in offering high aggregate bandwidth while delivering cost driven performance that is fully backward compatible with Ethernet, IP, and TCP. It features fault tolerance since it does not have single point of failure and its source routing protocol performs near shortestpath routing.





In [11] BCDC network depicted in Figure 23 was tolerance, and node-disjoint paths of BCDC using designed as server-centric data center network feasible experimental testbed. Though their using crossed cube model. The system is work highlighted superiority over DCell and Fatdecentralized while having recursively defined Tree implementation, fault-tolerant routing, and features needed to solve bandwidth bottleneck node-disjoint paths BCDC performance did not issues found at the upstream layered switch at show significant improvements over DCell and the tree structure for scalability. The design Fat Tree. The major issue is the inability of the ensured that the degree of BCDC server work highlight to scalability overhead instantiation minimized inter-connection cost. management especially in smart grid The work analysed data communication, fault applications.

www.idosr.org



Figure 23: BCDC Architecture [10] Generalized Hypercube and HSDC Architecture

Figure 23 reveals a low-cost interconnection architecture that relies on a generic hypercube, while for greater incremental scalability, Figure 25 portrays a highly scalable data center network architecture. This design would require complex reductions to make it lightweight for smart grid applications, despite the current low-cost interconnection architecture called the Exchanged Generalized Hypercube (EGH) and the High Scalability DCN architecture.



Figure 24: Generalized Hypercube Architecture [11]



Figure 25: HSDC Architecture **Stellar DCN Architecture**

In this section, the authors discussed a new generic construction for DCNs. The work presented a method of transforming the interconnection networks into potential dual-port server-centric DCNs. A description of the networking properties

So far, this work has reviewed the generic DCNs that lacks the capacity to run microservices in Cloud at scale. Big data center analytics using open-source tools like for stellar DCN such as routing algorithms, interconnection network, and the stellar transformation was discussed. The key aspects of the stellar construction are its topological simplicity, and its universal applicability.

Summary of Related Works

Hadoop, Apache Storm, Kafka and NoSQL Cassandra was discussed in [6]. The authors [13], presented a packet-switched optical network (PSON) architecture with

centralized control for intra-data-center connectivity. Though the work classified traffic flows with different characteristics. light weighted computation was not addressed. This is because, scalability, low-latency, high-speed, and energyefficient data center network remains verv critical for BDSMC deployment in future large-scale data centers. In [20], the authors focused on hybrid-stream big data analytics model for to performing multimedia big data analysis. Their model contains four procedures, i.e., data preprocessing, data classification, data recognition and data load reduction. In Multi-dimensional this regard, Convolution Neural Network (CNN) was proposed and evaluated. In [21], DC traffic classification was investigated for timesensitive and data-intensive service platforms. In [22], the authors proposed an edge processing unit that comprises two main parts: data classification model that classifies IoT data into maintenancecritical data (MCD) and maintenance-nondata and critical (MNCD) а data transmission unit that, based on the class of data. employs appropriate communication methods to transmit data to railway control centers. In [24], the work looked at the accurate estimation of data center resource utilization for multi-tenant co-hosted applications having dynamic and time-varying workloads. Their model adaptively and automatically identifies the most appropriate model to estimate DC resource utilization. The work in [25] proposed FlowSeer as a fast, low-overhead elephant flow detection and scheduling system using data stream mining. Their major idea is to leverage packets flows to

The development of DC network for efficient data stream offloading and Microservices in Cloud Computing Environments entails lots of efforts and Caps in literature

So far, existing works on DC network for efficient data stream offloading and Microservices orchestration within the cloud environments has some gaps. As a result of data collection with IoTs, stream computing, storage and DCN topologies

train the streaming classification models for accurately and quick prediction the rate and duration of any initiated flow. In the work proposed an online [27]. parameter-tuning method for the energyefficient DCN named high-speed optical layer 1 switch system for time-slotswitching-based optical DCNs). The DCN comprises optical circuit switching network, optical slot switching network, and electrical packet switching network for the spine layer. Also, a procedure for reconfiguring flow classification function and a method for online parameter tuning classification was discussed. In [12], the work discussed the processing of big data streams generated by the Industrial Internet of Things (IIoT). It also looked at the edge and cloud DCs. The work then introduced the Network Elephants Learner and anaLyzer (NELLY) as a novel and efficient method for applying incremental learning at the server side of software-defined data center networks (SDDCNs). The idea is to accurately and timely identify elephant flows with low traffic overhead. The work in [12] proposed autoencoder (AE) network based the distribution of polarimetric on synthetic aperture radar (POL-SAR) data matrix, called a mixture autoencoder (MAE). Through a detailed analysis of the data distribution POL-SAR data matrix. a normalization method was presented. In terms of micro-services in the cloud domain, little efforts has been made as well. For instance, a representee sample of literature were captured in FacGraph Micro-service architecture in [22] Computation offloading in [7]. Containerisation in [3].

CONCLUSION

contributions from data streams generation, to transmission, and storage for orchestration in the cloud.

Gaps in literature for further study

present conspicuous gaps as observed from reviewed literature as follows:

i. Absence of optimal Internet of Things (IoT) node partitioning in Clusters thereby affecting Quality of Service (QoS) management.

- ii. Network collapse resulting from point of failure topological layouts which affects workload scalability.
- iii. Partial service accessibility as well as non-availability of recursive sustained chain for traffic propagation.
- iv. Absence of Bayesian machine learning technique for QoS metric in BDMSC systems.
- Existing works have not developed v. optimization schemes for data in motion (BDMSC).
- Little work has been done on vi. connection availability model for stream generation
 - REFERENCES 1. Emara and J. Z. Huang,(2020). Distributed Data Strategies to Support Large-Scale Data Analysis Across Geo-Distributed Data Centers. IEEE Access, vol. 8, pp. 178526-178538. DOI:

10.1109/ACCESS.2020.3027675. 2. Zhong W., Yu N. and Ai C., (2020).

Applying big data based deep learning system to intrusion detection, Big Data Mining and Analytics, vol. 3, no. 3, pp. 181-195, DOI:

10.26599/BDMA.2020.9020003.

3. Zhang, Y. Zhang, L. Yin, T. Yuan, Z. Wu and H. Luo, (2019). Mining Frequent Items Over the Distributed Hierarchical Continuous Weighted Data Streams in Internet of Things, IEEE Access, vol. 7, pp. 74890-74898, DOI:

10.1109/ACCESS.2019.2911573.

- 4. Algahtani J. and Hamdaoui B. "Rethinking (2018).Fat-Tree Topology Design for Cloud Data Centers,", IEEE Global *Communications* Conference (GLOBECOM), Abu Dhabi, United 1-6, Arab Emirates, pp. doi: 10.1109/GLOCOM.2018.8647774.
- 5. Guo C., Lu G., Li D., Wu H., Zhang X., Shi Y., Tian C., Zhang Y., and Lu S.,(2009). "BCube: а high performance. server-centric network architecture for modular data centers," in Proc. of the ACM

- vii. Most works have not explored the collision domain mapping of IoT clustered subnet work for QoS optimization.
- viii. Most works lack optimal workload coordination for reduced resource drain.
- ix. Existing approaches offers heavy traffic overhead, lower scalability, lower accuracy, and high detection time.
- High performance Computing (HPC) x. workload managers lack microservices support and deeply integrated container management, as opposed to container orchestrators (e.g. Kubernetes).

SIGCOMM Conf. on Data Comm., pp. 63-74.

- 6. Guo D., Chen T., Li D., Li M., Liu Y, and Chen G.,(2013). "Expandable and cost-effective network structures for data centers using dual-port servers," IEEE Trans. on *Computers*, 62(7), Pp. 1303-1317.
- 7. Liao Y., Yin D., and Gao L., (2010). "Dpillar: Scalable dual-port server interconnection for data center networks," in Proc. of 19th Int'l. Conf. on Computer Comm. and *Networks*, pp.1-6.
- 8. Liu K, Lu Z. and Yan, J. (2018). "Parallel Wavelength Fault Tolerant Clos-network for Space Optical Network," 10th Int'l Conf., on Wireless Communications and Signal Processing (WCSP), Hangzhou, pp. 1-5, doi: 10.1109/WCSP.2018.8555941.

- 9. Radhakrishnan, M. Tewari, R. Kapoor, G. Porter and A. Vahdat, (2013). Dahu: Commodity switches for direct connect data center networks. Architectures for Networking and Communications Systems, San Jose, CA, pp. 59-70, DOI: 10.1109/ANCS.2013.6665176.
- 10. G. M., S. Raviteja, S. S. and R. Mahalakshmi, (2019). "Data Acquisition For Residential Energy Management Employing IoT Using ThingSpeak," IEEE Region 10 Symposium (TENSYMP),

Akawuku et al

Kolkata,India. pp. 272-276, Doi 10.1109/TENSYMP46218.2019.897 1366.

- 11. Haggag, M. M. Tantawy and M. M. S. El-Soudani,(2020). Implementing a Deep Learning Model for Intrusion Detection on Apache Spark Platform. IEEE Access, vol. 8, pp. 163660-163672, 2020, DOI: 10.1109/ACCESS.2020.3019931.
- 12. Liu K., Lu Z. and Yan J., (2018). Parallel Wavelength Fault Tolerant Clos-network for Space Optical Network, 10th Int'l *Conf., on Wireless Communications and* Signal Processing (WCSP), Hangzhou, pp. 1-5, DOI:

10.1109/WCSP.2018.8555941.

- 13. Lin Wang; Xinbo Wang; Massimo Tornatore; Kwang Joon Kim; Sun Me Kim: Dae-Ub Kim; Kyeong-Eun Han: Biswanath Muk, (2018)."Scheduling with machine-learningbased flow detection for packetswitched optical data center networks," in IEEE/OSA Journal of Communications and Optical Networking, vol. 10, no. 4, pp. 365-375, doi: 10.1364/JOCN.10.000365.
- 14. Alejandro E., Iain A., Stewart A., Javier N., Abbas E., Kiasari B., (2017). The stellar transformation: From interconnection networks to datacenter networks. DOI: 10.1016/j.comnet.2016.12.001
- 15. Al-Fares M., Loukissas A. and Vahdat A. (2008). "A Scalable Commodity Data Center Network Architecture" *Proc. of ACM SIGCOMM'08*, pp. 63-74.
- 16. Ankit Singla, Chi-Yao Hong, Lucian Popa, P. Brighten Godfrey, (2012). Jellyfish: Networking Data Centers Randomly", *arXiv:1110.1687v3* [cs.NI].
- 17. Baccarelli, N. Cordeschi, A. Mei, M. Panella, M. Shojafar and J. Stefa, (2016). Energy-efficient dynamic traffic offloading and reconfiguration of networked datacenters for big data stream mobile computing: review, challenges, and a case study. IEEE

Akawuku et al

Network, vol. 30, no. 2, pp. 54-61, DOI: 10.1109/MNET.2016.7437025.

- 18. Baig, W. Iqbal, J. L. Berral, A. Erradi and D. Carrera (2019). Adaptive Prediction Models for Data Center Resources Utilization Estimation, *IEEE Transactions on Network and Service Management*, 16(4), pp. 1681-1693, DOI: 10.1109/TNSM.2019.2932840.
- 19. Bagaa, T. Taleb, J. Bernal Bernabe and A. Skarmeta, (2020). QoS and Resource-aware Security Orchestration and Life Cycle Management. IEEE Transactions on Mobile Computing, DOI: 10.1109/TMC.2020.3046968.
- 20. Chao, K. C. Lin and M. Chen. (2019). Flow Classification for Software-Defined Data Centers Using Stream Mining," in *IEEE Transactions on Services Computing*, 12(1), pp. 105-116, DOI: 10.1109/TSC.2016.2597846.
- 21. Cicconetti C., Conti M., Passarella A. and Sabella D., (2020)."Toward Distributed Computing Environments with Serverless Solutions in Edge Systems," in IEEE Communications Magazine, 58(3), Pp. 40-46, doi: 10.1109/MCOM.001.1900498.
- 22. Chen, L. Lin, B. Li, Q. Wang and Q. Zhang,(2021). Silhouette: Efficient Cloud Configuration Exploration for Large-Scale Analytics. IEEE Transactions on Parallel and Distributed Systems, DOI: 10.1109/TPDS.2021.3058165.
- 23. Chkirbene, R. Hadjidj, S. Foufou and R. Hamila, (2020). LaScaDa: A Novel Scalable Topology for Data Center Network, *IEEE/ACM Transactions on Networking*, DOI: 10.1109/TNET.2020.3008512.
- 24. Chuanxiong Guo, Haitao Wu, Kun Tan, Lei Shi, Yongguang Zhang, Songwu Lu, (2008). "DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers", SIGCOMM'08, Seattle, Washington, USA.
- 25. Dolev P. Florissi E. Gudes S. Sharma and I. Singer (2019). A survey on

geographically distributed big-data processing using MapReduce. IEEE Trans. Big Data vol. 5 no. 1 pp. 60-80.

- 26. Domke, Satoshi, Matsuoka, Ivan Radanov , Yuki Tsushima, Tomoya Yuki. Akihiro Nomura. Shin'ichi Miura (2019). The First Supercomputer with HyperX Topology: A Viable Alternative to Fat-Trees. IEEE Symposium on High-Performance Interconnects (HOTI), Santa Clara, CA, USA, pp. 1-4, DOI: 10.1109/HOTI.2019.00013.
- 27. Estrada-Solano, O. M. Caicedo and N. L. S. Da Fonseca, (2020). NELLY: Flow Detection Using Incremental Learning at the Server Side of SDN-Based Data Centers. *IEEE Transactions on Industrial Informatics*, 16(2), Pp. 1362-1372, **DOI**: 10.1109/TII.2019.2947291.
- 28. Fujiwara I., Koibuchi M., Matsutani H. and Casanova H., (2014). Skywalk: A Topology for HPC Networks with Low-Delay Switches," 2014 IEEE 28th International Parallel and Distributed Processing Symposium, Phoenix, AZ, pp. 263-272, DOI: 10.1109/IPDPS.2014.37
- 29. Fu, G., Sun J. and Zhao J., (2018). An optimized control access mechanism based on micro-service architecture. 2018 2nd *IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing, China, pp. 1-5, DOI: 10.1109/EI2.2018.8582628.
- Gliksberg J., A. Capra, A. Louvet, P. J. García and D. Sohier (2019). High-Quality Fault-Resiliency in Fat-Tree Networks (Extended Abstract). 2019 *IEEE Symposium on High-Performance Interconnects (HOTI)*, Santa Clara, CA, USA, pp. 9-12, DOI: 10.1109/HOTI.2019.00015.
- 31. Guo Z. and Yang Y., (2014). Collaborative Network Configuration in Hybrid Electrical/Optical Data Center Networks. 2014 IEEE 28th Parallel International and Distributed Processing Symposium,

Akawuku et al

Phoenix, AZ, pp. 852-861, DOI: 10.1109/IPDPS.2014.92.

- 32. Mahmud, J. Z. Huang, S. Salloum, T. Z. Emara and K. Sadatdiynov, (2020).А survev of data partitioning and sampling methods to support big data analysis, in Big Data Mining and Analytics, vol. 3, no. 2, pp. 85-101, DOI: 10.26599/BDMA.2019.9020015.
- 33. Mao J., Han B., Lv G., Sun Z. and Lu X., (2015). FRINGE: Improving the scalability of Ethernet DCN via efficient software-defined edge control, *IEEE 23rd International Symposium on Quality of Service (IWQoS)*, Portland, pp. 63-64, DOI: 10.1109/IWQoS.2015.7404707.
- 34. Saki M., Abolhasan M. and Lipman J., (2020). A Novel Approach for Big Data Classification and Transportation in Rail Networks, *IEEE Transactions on Intelligent Transportation Systems*, 21(3), Pp. 1239-1249, DOI: 10.1109/TITS.2019.2905611.
- 35. Saber M. A. S., Ghorbani M., Bayati A., Nguyen K. and Cheriet M., (2020).Online Data Center Traffic Classification Based on Inter-Flow Correlations, *IEEE Access*, vol. 8, pp. 60401-60416, DOI: 10.1109/ACCESS.2020.2983605.
- 36. Jahanbakht, W. Xiang, L. Hanzo and M. R. Azghadi, (2021). Internet of Underwater Things and Big Marine Data Analytics—A Comprehensive Survey. IEEE Communications Surveys & Tutorials, DOI: 10.1109/COMST.2021.3053118.
- 37. Ji and Li B., (2016). Wide area analytics for geographically distributed datacenters. Tsinghua Sci. Technol. 21(2), pp. 125-135.
- 38. Jiang Z., Ding Z., Gao X. and Chen G., (2014). "DCP: An efficient and distributed data center cache protocol with Fat-Tree topology," The 16th Asia-Pacific Network Operations and Management Symposium, Hsinchu, pp. 1-4, doi: 10.1109/APNOMS.2014.6996589.
- 39. Jiang W., Qi J., X., J. Yu, J. Huang and R. Zhang, (2019). "HyperX: A

Scalable Hypergraph Framework," in IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 5, pp. 909-922, doi: 10.1109/TKDE.2018.2848257.

- 40. John Kim, William J. and Dennis Abts, (2007). "Flattened Butterfly: A Cost-Efficient Topology for High-Radix Networks", *ISCA'07*, San Diego, California, USA.
- 41. Kan S., Fan J., Cheng B. and X. Wang, (2020). "The Communication Performance of BCDC Data Center Network "12th International Conference on Communication Software and Networks (ICCSN), Chongqing, China, pp. 51-57,doi: 10.1109/ICCSN49894.2020.913907 7.
- 42. Ramasamy, Deep Medhi Karthik (2018). Routing and Traffic Engineering in Data Center Networks. Network Routing (Second edition), Algorithms, Protocols, and Architectures. the Moraan *Kaufmann Series in Networking*, Pp. 396-422.

https://doi.org/10.1016/B978-0-12-800737-2.00014-4

- 43. Mao J., Han B., Lv G., Sun Z. and Lu X., (2015). FRINGE: Improving the scalability of Ethernet DCN via efficient software-defined edge control, *IEEE 23rd International Symposium on Quality of Service (IWQoS*), Portland, pp. 63-64, DOI: 10.1109/IWQoS.2015.7404707.
- 44. Murakami M., Kubokawa H., Sugiura Oki E., Okamoto S. K., and Yamanaka N., (2020). Online parameter tuning of the flow classification method in the energyefficient data center network "HOLST". IEEE/OSA J. of Optical Communications & Networking. 12(11).344-354. DOI: Pp. 10.1364/JOCN.395798.
- 45. Udeze C. C., Okafor K. C., Okezie C. C., Okeke I. O., Ezekwe G.C., (2014). "Performance Analysis of R-DCN Architecture for Next Generation Web Application Integration", In IEEE Xplore Digital Library, 6th IEEE Int'l Conf., on

Akawuku *et al*

Adaptive Sci., & Tech., (ICAST 2014), Covenant University Otta, Nigeria .Pp.1-12.

- 46. Qiao P., X. Wang, X. Yang, Y. Fan and Z. Lan, (2017).Preliminary Interference Study About Job Placement and Routing Algorithms in the Fat-Tree Topology for HPC Applications. *IEEE International Conference on Cluster Computing (CLUSTER)*, Honolulu, HI, 2017, pp. 641-642, DOI: 10.1109/CLUSTER.2017.90.
- 47. Xiaohui Ye, P. Mejia, Yawei Yin, R. Proietti, S. J. B. Yoo and V. Akella, (2010). DOS - A scalable optical switch for datacenters, *ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*, La Jolla, CA, 2010, pp. 1-12.
- 48. Xia Y., Hamdi M. and Chao H. J., (2016). A Practical Large-Capacity Three-Stage Buffered Clos-Network Switch Architecture, *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 317-328, DOI: 10.1109/TPDS.2015.2408614.
- 49. Xu C., Wang K., Sun Y., Guo S. and Zomaya A. Y., (2020). Redundancy Avoidance for Big Data in Data Centers: A Conventional Neural Network Approach, *IEEE Transactions on Network Science and Engineering*, 7(1), pp. 104-114, 1, DOI:

10.1109/TNSE.2018.2843326.

- 50. Xiaohui Ye, P. Mejia, Yawei Yin, R. Proietti, S. J. B. Yoo and V. Akella, (2010). DOS - A scalable optical switch for datacenters, ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS), La Jolla, CA, pp. 1-12.
- 51. Van-Dai Ta, Chuan-Ming Liu and G. W. Nkabinde, (2016). Big data stream computing in healthcare real-time analytics. *IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, China, pp. 37-

www.idosr.org

42, 10.1109/ICCCBDA.2016.7529531.

- 52. Wang J., Hou B., L. Jiao and S. Wang, (2010).POL-SAR Image Classification Based on Modified Stacked Autoencoder Network and Data Distribution, IEEE Transactions on Geoscience & Remote Sensing, 1678-1695, 58(3), Pp. DOI: 10.1109/TGRS.2019.2947633.
- 53. Wu, (2020).Cloud-Edge Y. Orchestration for the Internet-of-Architecture and Things: AI-Powered Data Processing. IEEE Internet of Things Journal, DOI: 10.1109/JIOT.2020.3014845.
- 54. Wang G., Lin C., J. Fan, B. Cheng and X. Jia, (2020). "A Novel Low-Cost Interconnection Architecture Based on the Generalized Hypercube," in IEEE Transactions on Parallel and Distributed Systems, vol. 31, no. 3, 647-662, doi: pp. 10.1109/TPDS.2019.2941207.
- 55. Zhai, Y. Yang, H. Wang and S. Du, (2020).Multi-attention fusion modeling for sentiment analysis of educational big data, Big Data Mining and Analytics, vol. 3, no. 4, 311-319, DOI: pp. 10.26599/BDMA.2020.9020024.
- 56. Zhou N., Georgiou Y., Zhong L., Zhou H. and Pospieszny M., (2020). "Container Orchestration on HPC IEEE Systems," 2020 13th International Conference on Cloud Computing (CLOUD), Beijing, China, doi: 34-36. pp. 10.1109/CLOUD49709.2020.00017.