International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 106-113, August-2023
https://doi.org/10.47001/IRJIET/2023.708014

Simulation of Big Data Stream Mobile Computing Architecture (BDSMCA) Data Center Network (DCN) for Efficient Data Stream Offloading in Cloud Environments

[1]Godspower I. Akawuku, [2]Anusiuba, Overcomer Ifeanyi Alex, [3]Paul U. Roseline, [4]Samauel O. Adejumo

[1,2,3,4]Nnamdi Azikiwe Unversity, Awka (NAU), Nigeria
Authors E-mail: [1]gi.akawuku@unizik.edu.ng, [2]oi.anusiuba@unizik.edu.ng, [3]ru.paul@unizik.edu.ng

Abstract **- Simulation of a Big data streams data center (BDSDC) reengineered for efficient dumpsite datastream offloading. A use case municipal waste management aggregation data center network (DCN) and cloud-driven micro-services orchestration at the edge with low latency translation into the cloud environments. Discrete-Event Modelling and Simulation Methodology (DEMSM) was adopted. Using the experimental test data gathered from the experimental testbed (UNN DCN), a simulation study was carried out in Riverbed Modeller while allowing for result comparison with the trace file of the typical traditional DCN. It was discovered that BDMSC performed much better that the traditional DCN and addressed majority of the challenges. The results of the proposed BDSDC system considered BDSCA Optimization, and non-BDSCA Optimization use-cases. Second, the proposed BDSA was then compared with Bayesian and MapReduce algorithms. With BDSCA Optimization, 47.37% data stream workload is provisioned which is very useful in deterministic traffic workloads. This is in contrast with the best-efforts scheme that yielded 52.63%. In terms of throughput, proposed BDSA offered 52.63% throughput cycles while Bayesian and MapReduce gave 36.84% and 10.53% each. Considering network latency, the proposed BDSCA latency optimization is shown to be very attractive at 27.77%. This is certainly better than Bayesian (55.56%) and MapReduce (16.67%). In terms of resource utilization, at peak traffic, all the algorithms had similar trend pattern even at the steady and relaxed states. At a closer experimental control and monitoring, the resource utilization, MapReduce Apriori, Bayesian and the proposed BDSCA offered 37.55%, 25.03% and 37.42% respectively. Finally, this is better than reactive DCell and BCube integration cloud domains.**

Keywords: Datacenter, Datastream, Offloading, simulation, Cloud Environment.

1. Introduction

Big data streams data center (BDSDC) is a computational server farm or a compute structure where majority of volumetric input and output streams servers as well as storage systems are located, operated, and managed using structured pipeline methodology. It is also referred to as the consolidation point for provisioning multiple services that drive big real time business processes. In BDS-DC designs, the emphasis on low latency, real-time data dissemination for data flows in a highly distributed, available coordination service domain. Data streaming continuously generates data from different sources. This now creates a new taxonomy known as Big Data Stream Mobile Computing (BDSMC). Examples of streaming services within BDSMC context include edge/ Internet of things (IoT) traffic, social media traffic, Netflix, iTunes, Hulu, YouTube, Vudu and Amazon Instant. These types of traffic are greatly impacting the quality of service found in the Cloud data centres. Incidentally, today's Cloud Computing domain is primarily based on proprietary datacentres, where hundreds of thousands of dedicated servers are setup to host the cloud services asserted Mengistu *et al.* (2017).

A re-engineered simulation system for Big Data Stream Data center (BDSDC) network for efficient data stream offloading and micro-scaled services in cloud driven computing environment was developed. BDSDC is introduced as a waste management cloud aggregation (WMCA) construct bounded to modern cities with low computational overhead. It solves the problems of non-existent edge-to-cloud data center networks for end-to-end service provisioning of big data streams from disaggregated dumpsites. The aim of this research was to develop a simulation re-engineered DC network for efficient data stream offloading in cloud computing environments. The Objectives were to carry out Experimental studies on typical test beds to ascertain the limitations of the traditional data stream offloading (DO) models, compare the performance of the proposed reengineered datastream offloading DCN model with the existing DCN architectures using real-time data from a typical DCN.

2. Literature Survey

Clearly, the traditional Data Center Networks (DCNs) are suffering from many problems including high energy consumption, high latency, fixed throughput of links and limited reconfigurability to the traffic demand Fayyaz and Aziz (2014). The classification levels of data centers represent the design certification. A tier depicts the service level and there are 4 tiers of Data centers are Obayi and Okafor (2020) submissions: Tier 1 Data Center, Tier 2 Data Center, Tier 3 Data Center, Tier 4 Data Center.

A tier 1 data center can be little more than a powered warehouse and it is not complex in structure, hence unreliable for stream computing. Tier 4 offer guarantee of uptime and 2N (two times the amount required for operation) cooling and redundant power and infrastructure. Level 4 edge devices does not have issues at the data center infrastructures due to its redundancies. A tier 3 data center can perform repairs without noticeable service disruption. A level 3 provider offer an N+1 (the amount required for operation plus a backup) availability for clients. level 3 is even tolerant of some faults. Tier 4 data centers are considered “fault tolerant.” Unplanned maintenance does not stop the flow of data to a data center Tier 4. Day-to-day operations continue regardless of any support taking place. Availability according to data center Tiers is highlighted below:

- Tier 1 – 99.671% Guaranteed availability
- Tier 2 – 99.741% Guaranteed availability
- Tier 3 – 99.982% Guaranteed availability
- Tier 4 – 99.995% Guaranteed availability

At Tier 4, a data center must adhere to the following:

Zero single points of failure. Tier 4 providers have redundancies for every process and data protection stream. No single outage or error can shut down the system.

- 99.995% uptime per annum. This is the level with the highest guaranteed uptime. It must be maintained for a center to maintain Tier 4 ranking.
- 2N+1 infrastructure (two times the amount required for operation plus a backup). 2N+1 is another way of saying “fully redundant.”
- No more than 26.3 minutes of downtime per annum as a maximum figure. Providers must allow for some downtime for optimized mechanical operations; however, this annual downtime does not affect customer-facing operations.
- 96-hour power outage protection. A level 4 infrastructure must have at least 96 hours of independent power to qualify at this tier. This power must not be connected to any outside source and is entirely proprietary. Some centers may have more.

At Tier 3, a data center must adhere to the following:

1) N+1 (the amount required for operation plus a backup) fault tolerance. A Tier III provider can undergo routine maintenance without a downtime in operations. Unplanned maintenance and emergencies may cause problems that affect the system. Problems may potentially affect customer-facing operations.
2) 72 hours of protection from power outages. This provider must have at least three days of exclusive power. This power cannot connect to any outside source.
3) No more than 1.6 hours of downtime per annum. This downtime is allowed for purposes of maintenance and overwhelming emergency issues.
4) 99.982 % uptime. This is the minimum amount of uptime that a level 3 provider can produce. The redundancies help to protect this number even if a system suffers unexpected issues.

The physical network topology of most traditional data center networks is typically organized as a three-layer hierarchy, as shown in Figure 2.1. The layers include: the access layer, the aggregation layer and the core layer.

The Core Layer: The core layer provides the high-speed packet switching backplane for all the flows that is going in and out of the data center. The core layer provides connectivity to multiple aggregation modules and provides a resilient Layer 3 routed fabric with no single point of failure. The core layer runs an interior routing protocol, such as OSPF or EIGRP, and load balances traffic between the campus core and aggregation layers. Typically, the core layer utilizes high performance low latency switches providing high densities of 10GE which is used to link up to the aggregation layer switches. Switches at this layer operate exclusively as Layer 3 devices.

The Aggregation Layer: The aggregation layer modules provide important functions, such as service module integration, Layer 2 domain definitions, spanning tree processing, and default gateway redundancy. Server-to-server multi-tier traffic flows through the aggregation layer and can use services, such as firewall and server load balancing, to optimize and secure applications. These modules provide services, such as content switching, firewall, secured service layer (SSL) offload, intrusion detection, network analysis, and more. In other words, we can say that aggregation layer acts as a service layer for the data center network. Multiple access layer switches will also use the aggregation layer as an interconnection point. The use of 10GE links to uplink into the Core Layer is a common practice. More of an emerging trend is the use of 10GE links to downlink into the access layer providing higher bandwidth and future proofing the network.

The Access Layer: The access layer is where the servers physically attach to the network. The server components consist of 1 Rack Unit (RU) servers, blade servers with integral switches, blade servers with pass-through cabling, clustered servers, and mainframes with Operating System Adapters (OSA).

The access layer network infrastructure consists of modular switches, fixed configuration 1 or 2RU switches, and integral blade server switches. Switches provide both Layer 2 and Layer 3 topologies, fulfilling the various server broadcast domain or administrative requirements. Hence the access layer provides connectivity for the many servers that deliver application and web services to the business as well as the interconnections for a server cluster design. An access Layer switches may be required to support both single and dual homed servers.

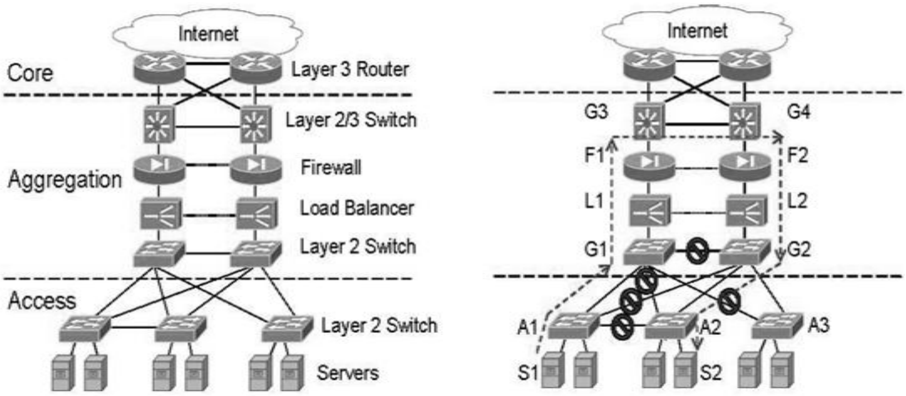


Figure 2.1: A 3-layer DCN topology including Firewalls (Ba-hutair et al., 2021)

Figure 2.1 is a typical three-layer DCN model and is comprised of access switches, aggregation switches and core switches. The access switches were connected to the data center servers such as database servers, application servers and web servers. The aggregation switches were used for the aggregation of the access switches while the core switches provide routing to and from the enterprise core network. In designing a data center network using three-tier design model, the main consideration is scalability since it is based on hierarchical design. Hence new aggregation switch pairs could be added with no need to modify the existing aggregation pairs. In this model, full mesh is not required as data center core switches do the routing.

The three-tier design model has some disadvantages which include:

- Higher latency due to additional layer.
- Additional congestion and oversubscription in the design.
- More managed nodes adding a certain amount of complexity for operation and maintenance.
- Higher energy consumption.
- The need for additional space.

In the two-layer data center model, the aggregation layer is collapsed into the core layer based on port density, aggregation throughput and oversubscription requirements as shown in figure 2.2. Here the access switches that are used for server connectivity are collapsed in high density core/aggregation switches. These switches provide the switching and routing functionalities for access switching interconnections and the various server virtual local area networks (VLANs). This implies that aggregation of access switches are done at the core layer instead of the normal aggregation layer as in the case of the three-layer design model. This aggregation of the access switches at the core layer allows for more flexibility and easier support for virtualization except that it requires very high-speed processing and high availability (HA) levels. Advantages of two-tier design model include: The number of devices used is dramatically reduced and that offers significant power savings, reduces the facilities footprint of the system, offers simplified device management, reduces the number of system failure points, and allows tighter security control.

- Design simplicity due to fewer switches and so fewer managed node.
- Reduced network latency since the number of switch hops has been reduced.
- Reduced network design oversubscription ratio.

However, for the two-layer DCN model, the connection between aggregation switch pairs must be fully meshed with high bandwidth so as to avoid bottlenecks in the network design. An aggregation switch pair is running routing protocols; hence, more switch pairs mean more routing protocol peering, more routing interfaces and complexity introduced by a full mesh design. In this work, Data center Network (DCN) reengineering we adopted is the two-layer data center model as shown in figure 2.2 due to its advantages over the three-layer model.

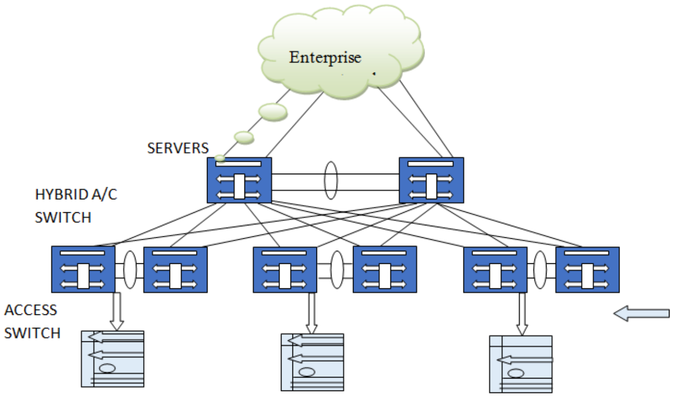


Figure 2.2: Two-Layer Data Center Model

Vendors Use Case – IBM Cloud Environment for Recharacterization and Design.

Offensive techniques/Data stream offloading - All elements are Hyper intelligent/reactive/resilience. Sensors are embedded intelligence, no congestion at this point (every system is having real time updates), Resilient, Adaptive.

Defensive techniques/Convention traffic Engineering - Load management/ auto-scaling/ buffers / manual for QoS management. Server will wait and monitor traffic workload. Buffer overflows at the cloud gateway drops packets creating zero connection and leads to downtime. When managed by an application that supports virtual machine failover, applications can remain highly available and tolerant in the event of any physical server failure.

Hence, even if the physical host server system fails, this will not affect the application availability. In these instances, the introduction of virtual machine failover in the DCCN Infrastructure of the smart green energy management system can offer significant benefits at large.

3. Problem Definition

Currently, there are non-existent edge-to-cloud datacenter networks for end-to-end service provisioning of big data streams such as Big Data Stream Mobile Computing Data center (BDSMC). Today, streams of big data are post-processed and stored in the legacy Data centers DCs. Virtualization has made it possible thereby allowing synchronously execution on the virtual machines (VM), hosted by the DC. For efficient data stream and offloading, low latency computational time with reduction in device deployment has not been addressed. Also, Quality of Service (QoS) issues for BDSMC in most DCs have not been investigated. DCs implements massive workload parallel processing in clusters which in turn requires large amounts of data to be transferred among different VMs. There is need to improve the overall processing time, especially for stream workloads.

One common feature of BDMSC is the real-time nature resulting from the need to serve users in a timely fashion. Consequently, a new re-engineering is required for datacenter network operation handling stream workload. Non-availability of edge computing and ultra-reliable low-latency communication (URLLC) is yet to be addressed for edge-to-cloud domain.

Consider a real-life scenario in a smart factory floor where edge devices are ready to execute transactions. Delayed or inaccessible network resulting downtime network and servers will certainly affect continuity operations. Server downtime and poor network performance in data streamed traffic workload can affect the network efficiency.

A poor design of the datastream DCN where their data are stored, operated and managed is usually the root cue. The designers of such DCN are yet to consider the design goals for data centers vis-à-vis future expansion: scalability, high bandwidth and fault-tolerance. Consequently, network congestion results as the number of customer’s increases and integration of micro-service application becomes difficult since service provisioning and analytics were not originally envisaged for in the design. This work on the development of data center network for efficient datastream offloading and cloud micro-services seeks to solve these problems using an efficient edge to cloud server interconnection scheme with a collision suppression algorithm. Hence, in this work, high network bandwidth, scalability and fault-tolerance are the design goals in order to solve this problem.

4. Methodology

Discrete-Event Modelling and Simulation Methodology (DEMSM) was adopted. Using the experimental test data gathered from the experimental testbed (UNN DCN), a simulation study was carried out in Riverbed Modeller while allowing for result comparison with the trace file of the typical traditional DCN. It was discovered that BDMSC performed much better that the traditional DCN and addressed majority of the challenges.

5. Simulation and Evaluation

For the simulation testbed in Figure 5.1, the following were introduced viz: OS Platform/NetC, Cloud Power-Edge C6420, Cisco WS-C3650-24TS-L 24 Port Switch, Cisco UCS C480 ML server core. The locations of edge nodes and data sources are as shown in Figure 3.1. The simulation configuration affects resource use per node and service characteristic. The computing resources available to each network node used as an edge node are limited. The network is supposed to have multiple types of services, each of which provides edge services by running as many APIs as the maximum number of edge nodes. The delay time of 2.057 (2 minutes and 1 second) is used for non-mission critical data recovery. This will reduce the load intensity of the network during the recovery of any lost data on average irrespective of location.

Figure 5.1 shows the edge testbed construction used for moving data streams from the dumpsite into the cloud. The generated API key by the server (thingspeak) has been previously highlighted. This API key is the optimized service used to push data to the server from the IoT sensor nodes. Also, high-level security end-to-end encryption was explored.

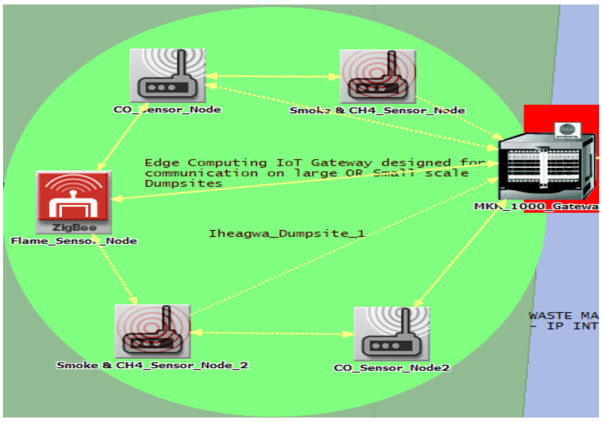


Figure 5.1: IoT dumpsite sensor field

Now, to quantify the performance of the selected API orchestration algorithms, the network experimental simulation procedure was carefully designed with Riverbed trace file engine. This took into consideration of the baseline metrics for the three scenarios. In the simulation environment, data stream measurement object palettes were used for the validation study. The objective of the dumpsite data stream simulation analysis justifies baseline metrics for API orchestration.

6. Validation of API Orchestration Algorithm

In this research, the use of API orchestration for BDSMC scenarios was investigated. To achieve Edge-to-Cloud communication, the BDSMC RESTful Orchestration algorithm was introduced in Figure 4.1 via the application and Configuration palettes of the Riverbed modeler. Other scenario instances were created in the Riverbed also. Using Riverbed modeler 17.8 with microservices plugins, the optimal stimulation parameters used for evaluating the established API orchestration algorithm are presented in Figure 6.1. Within a 100 x 100m2 dumpsite sensor field, five (5) edge IoT were distributed at random. Each sensor's transmission range was set to 30 meters, however, it can be increased to 100 meters. For minimum error deviation, the maximum communication channel bit rate was initially set to 300kbps. The size of each packet was set to 256bytes. The control packet size Request-to-Send (RTS) has been set to 2346 (max data stream frame size), clear to send (CTS), an acknowledge (ACK) to 5 bytes. The weight utilized in the weighted moving average distribution was set to 0.1, and the maximum active lifetime of IoT nodes was set at 0.5J. The IoT dumpsite sensor nodes in the simulation simulated the (MAC) layer protocol while sensing.

The node buffer size was changed following the data speeds from 10 to 250kbps for the best simulation response. The total queue length for the node was set to 10000 packets, with each queue size set to accommodate a maximum of 15000 packets. This experiment used a fixed workload consisting of 5 sources and 1 sink with logical pod instances throughout the experiment. The initial originating rate was 4pps, with a maximum originating rate of 16pps. The IEEE 802.11e MAC protocol provided in the Riverbed Network Simulator was used in this work. The parameters are selected to largely avoid hidden node and overhead issues during active transmission.

So far, this work has discussed extensively the processes of simulation execution. To further validate the API orchestration used in BDSMC, the Riverbed modeler library (Chen et al., 2020) was mapped with Cloud API keys from Thingspeak (Haggag et al., 2020). This facilitated the needed trace files and statistics associated with functional deployment. The results of the simulation are then presented. Service throughput, latency profiles, and auto-scaling were introduced in the experiment to test performance at full deployments. After the simulation run, data were collected from Figure 6.1 using seven nodes in different clusters.

This work used API orchestration services to validate the study. In this regard, this study carried out a scenario-based study using related optimization algorithms. Hence, this work compared the proposed BDSMC RESTful Orchestration with ATCloud (Zaharia, 2016), (Mahmud et al., 2020) and Karamel_API_Orchestration (Ji and Li, 2016). The edge module operates with an active battery from 12V solar panels deployed in the experimental testbed.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 106-113, August-2023
https://doi.org/10.47001/IRJIET/2023.708014

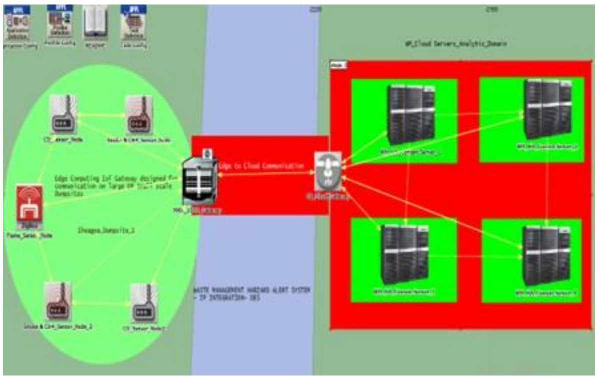


Figure 6.1: BDSMC Validation scenario for the proposed AO Algorithm

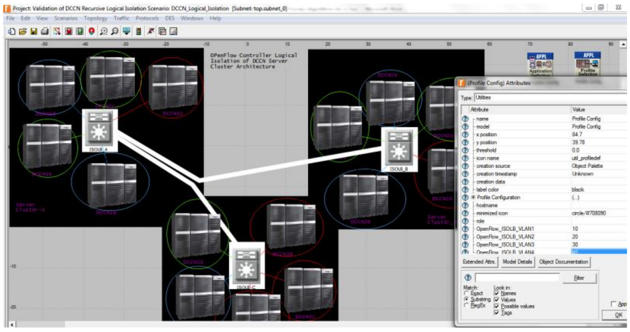


Figure 6.2: BDSMC OpenFlow Server Cluster logical Isolation Strategy with Riverbed Modeler, Academic Version 17.5

7. Datasets Generation & Plots Validation

In this section, the contribution to remote traffic communication in terms of QoS metrics of the various API orchestration algorithms (proposed BDSMC, ATCloud, and Karamel) is presented. At traffic workload, the network distributed rate adjustments of the queues on the dumpsite sensor nodes and cluster-head gateways. This is very important for end-to-end data streaming in Figure 4.14. The algorithm interfacing is simply done in the discrete event simulation engine using the object palette and configuration windows on a scenario basis. All the simulation objects were imported and seen as trace files once the compiler successfully executes the scenario algorithms/procedures. At this point, the simulation run-time generates the expected result plots concurrently.

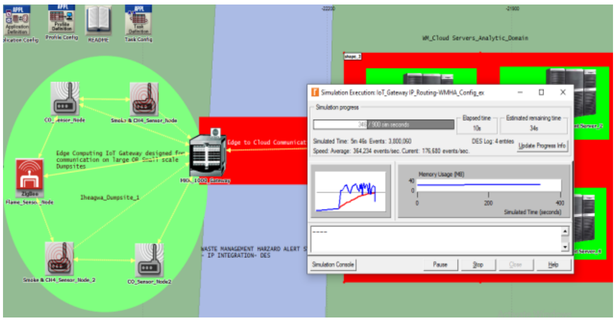
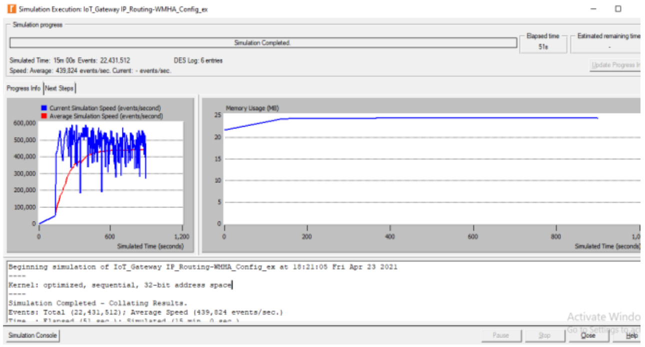


Figure 7.1: BDSMC Compilation process scenario AO Algorithm



Elapsed time: 51s
Estimated remaining time: -
Simulated Time: 15m 00s Events: 22,431,512
DES Log: 6 entries
Speed: Average: 439,824 events/sec. Current: - events/sec.

```
Beginning simulation of IoT_Gateway IP_Routing-WMHA_Config_ex at 18:21:05 Fri Apr 23 2021
----
Kernel: optimized, sequential, 32-bit address space
----
Simulation Completed - Collating Results.
Events: Total (22,431,512); Average Speed (439,824 events/sec.)
```

Figure 7.2: Successful validation scenario for the proposed AO Algorithm

8. Result and Discussion

Figure 8.1 shows that with BDSCA Optimization, 47.37% data stream workload is provisioned which is very useful in deterministic traffic workloads. This is in contrast with the best-efforts scheme that yielded 52.63%.

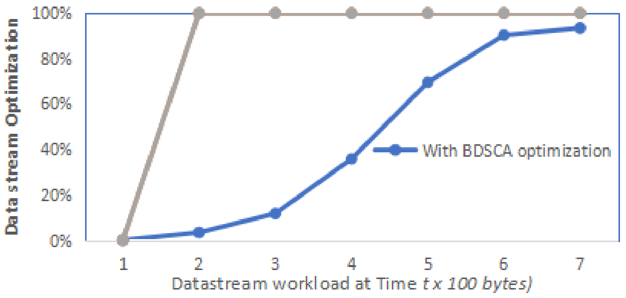


Figure 8.1: BDSCA Optimization validation

In figure 8.2, this is in contrast with the best-efforts scheme that yielded 52.63%. In terms of throughput, proposed BDSA offered 52.63% throughput cycles while Bayesian and MapReduce gave 36.84% and 10.53% each. Considering network latency, the proposed BDSCA latency optimization is shown to be very attractive at 27.77%.

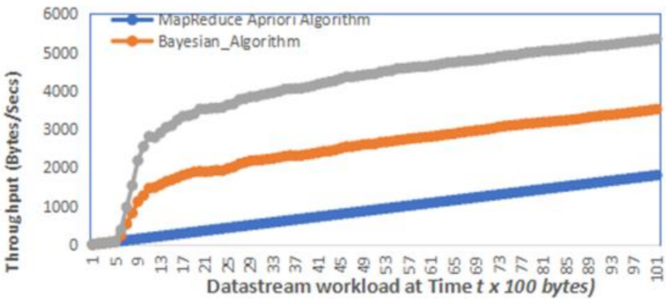


Figure 8.2: BDSCA Throughput Optimization

This is certainly better than Bayesian (55.56%) and MapReduce (16.67%). In terms of resource utilization, at peak traffic, all the algorithms had similar trend pattern even at the steady and relaxed states, shown in Figure 8.3.

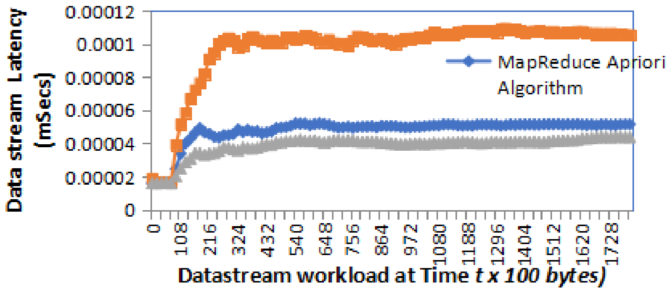


Figure 8.3: BDSCA latency Optimization

At a closer experimental control and monitoring, the resource utilization, MapReduce Apriori, Bayesian and the proposed BDSCA offered 37.55%, 25.03% and 37.42% respectively. As showed in Figure 8.4.

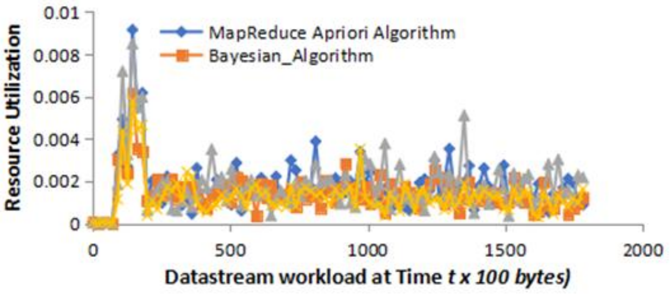


Figure 8.4: BDSCA Resource Optimization

9. Conclusion and future scope

In this paper, a re-engineered for BDSDC network for efficient data stream offloading and micro-scaled services in cloud driven computing environment is developed. BDSDC is introduced as a waste management cloud aggregation (WMCA) construct bounded to modern cities with low computational overhead. It solves the problems of non-existent edge-to-cloud data center networks for end-to-end service provisioning of big data streams from disaggregated dumpsites. The BDSDC depends on the WMCA concept for service provisioning dumpsite full duplex datastream efficiency. The system design life cycle was contextualized while using BDSCA to address computational complexities. The computational and architectural environment of BDSMC for WMCA is formulated using the joint probability density function. This employs a data splitter that handles large dataset splitting to obtain smaller data samples for data analytics. The QoS budgets and computational resources are represented in the BDMSC Computational System Architecture. The top level comprises the dumpsite location maps hosted by the sensor layers, the distributed Fog layer with OpenFlow load balancer.

Meanwhile, the DCN offers resources such as server processors, memories, storage applications, and networking infrastructure that are provisioned by the server services over the internet. One of the key components of DCN is resource virtualization which provides computing and storage services. A framework for container-based virtualization was achieved. This optimizes hardware usage through resource reuse and multiplexing which decreases the cost of power, hardware, and network bandwidth. For the proposed DCN, network input/output virtualization was used provide connectivity to the server virtual machines. Also, this computing system offers stability for dumpsite services aggregation. The implication is that multi-dumpsite parameters can be easily distributed for baseline analytics. In terms of throughput metrics, the use of large-scale computing architecture to actively support dumpsite process management within the cloud clusters was shown to be very significant. The proposed BDSA offered better throughput compared Bayesian and MapReduce schemes. The results demonstrate a brand-new cloud module for BDSMCaaS that is built on OpenStack API. In BDSMCaaS, the proposed BDSA was then used to validate the datastream provisioning. Considering BDSCA Latency Optimization, the proposed BDSCA latency optimization is shown to be very attractive when compared with Bayesian and MapReduce. The implication is that dumpsite workloads will be delivered a much faster rate allowing for quick analytics from the edge or the cloud. In terms of resource utilization, it was shown that at peak traffic, resource utilization was relatively provisioned by MapReduce Apriori, at the proposed BDSCA. This implies that Microservice-based architecture usage in MapReduce Apriori and proposed BDSCA will massively grow in acceptance when used in global system architecture. Using microservice containerization technique in BDSMC will speed up the deployment of dumpsite analytic applications since the network, CPU, memory, and storage performance will be running at optimum capacity.

Therefore, proposed BDMSC is efficient, scalable, cost effective, service-oriented, and responsive to dumpsite service provision, with rapid service delivery. As such, this thesis recommends the BDMSC-BDSCA for greater efficiency in Data stream offloading.

REFERENCES

[1] Ba-hutair, A. Bouguettaya and A. Ghari Neiat, (2021). “Multi-Perspective Trust Management Framework for Crowd sourced IoT Services”. *IEEE Transactions on Services Computing*, DOI: 10.1109/TSC.2021.3052219.

[2] Chen C., Hung L., Hsieh S., Buyya R. and Zomaya A. Y., (2020). "Heterogeneous Job Allocation Scheduler for Hadoop MapReduce Using Dynamic Grouping Integrated Neighboring Search," *in IEEE Transactions on Cloud Computing*, vol. 8, no. 1, pp. 193-206, doi: 10.1109/TCC.2017.2748586.

[3] Haggag M., Tantawy M. M. and El-Soudani M. M. S., (2020)."Implementing a Deep Learning Model for Intrusion Detection on Apache Spark Platform," *in IEEE Access*, vol. 8, pp. 163660-163672, doi: 10.1109/ACCESS.2020.3019931.

[4] Obayi A.A., and Okafor K.C., (2020). “Cloud-Fog Orchestration Infrastructure for Ante-natal Health Care Systems”. *2020 IEEE International Conference in Mathematics, Computer Engineering & Computer Science (ICMCECS).* Anchor University Lagos, Nigeria.

[5] Ji and Li B., (2016). “Wide area analytics for geographically distributed datacenters”. *Tsinghua Sci. Technol*. 21(2), pp. 125-135.

[6] Mengistu, A. Alahmadi, A. Albuali, Y. Alsenani and D. Che, (2017). "A" No Data Center Solution to Cloud Computing. *IEEE 10th International Conference on Cloud Computing (CLOUD), Honolulu, CA*, pp. 714-717.DOI: 10.1109/CLOUD.2017.99.

[7] Mahmud M. S, Huang J. Z., Salloum S., Emara T. Z. and Sadatdiynov K. (2020)."A survey of data partitioning and sampling methods to support big data analysis" *Big Data Mining Anal*. vol. 3 no. 2 pp. 85-101.

[8] Zaharia M. (2016). "Apache spark: A unified engine for big data processing" *Commun. ACM* vol. 59 no. 11 pp. 56-65.

Citation of this Article:

Godspower I. Akawuku, Anusiuba, Overcomer Ifeanyi Alex, Paul U. Roseline, Samauel O. Adejumo, “Simulation of Big Data Stream Mobile Computing Architecture (BDSMCA) Data Center Network (DCN) for Efficient Data Stream Offloading in Cloud Environments” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 106-113, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708014

****