

Understanding SDN architecture for addressing performance issues for OTT applications

Rounak R. Gupta

Research Scholar, University of Technology, Jaipur

Dr. Pramod Sharma

Professor, University of Technology, Jaipur

rounak_gupta@rediffmail.com

DECLARATION: I AS AN AUTHOR OF THIS PAPER / ARTICLE, HEREBY DECLARE THAT THE PAPER SUBMITTED BY ME FOR PUBLICATION IN THE JOURNAL IS COMPLETELY MY OWN GENUINE PAPER. IF ANY ISSUE REGARDING COPYRIGHT/PATENT/ OTHER REAL AUTHOR ARISES, THE PUBLISHER WILL NOT BE LEGALLY RESPONSIBLE. IF ANY OF SUCH MATTERS OCCUR PUBLISHER MAY REMOVE MY CONTENT FROM THE JOURNAL WEBSITE. FOR THE REASON OF CONTENT AMENDMENT/OR ANY TECHNICAL ISSUE WITH NO VISIBILITY ON WEBSITE/UPDATES, I HAVE RESUBMITTED THIS PAPER FOR THE PUBLICATION. FOR ANY PUBLICATION MATTERS OR ANY INFORMATION INTENTIONALLY HIDDEN BY ME OR OTHERWISE, I SHALL BE LEGALLY RESPONSIBLE. (COMPLETE DECLARATION OF THE AUTHOR AT THE LAST PAGE OF THIS PAPER/ARTICLE)

Abstract:

The expansion and need for new services provided by networks are the main forces behind research and development in this area. Guarantees of Quality of Service (QoS) and sufficient network capacity are crucial to satisfying this need. There are three tiers of service, or "applications," according to IEEE 802.1Q. These are Expedited Forwarding (EF), Assured Forwarding (AF), and Best Effort (BE). EF has the lowest loss, latency, and jitter and is not queued since it is the highest priority. Planning a Quality-of-Service policy allows you to set priorities for network traffic and services. The prioritized data can subsequently be assigned to a certain AF tier. Apps that use less bandwidth and are less sensitive, such as MULTIMEDIA, will be classified as AF, whereas Emails and P2P applications go under BE.

Keywords: *Quality, service, multimedia, e-mails, network, traffic, etc.*

1. INTRODUCTION

Test the controller's performance and the Open Flow switches in the testing network before implementing SDN in any network. We looked at the traffic characteristics of mobile networks with SDN support. To accurately describe the characteristics of the traffic, we create a flow-based internet traffic analysis model. Many middle boxes and forwarding devices are used in the construction of current mobile backhaul networks. Each gadget has intelligence built into it to carry out tasks. However, despite advancements in mobile network technology and a lack of available resources, mobile

service providers are unable to satisfy the growing customer demand for data and changing traffic patterns. Customers only sometimes utilize the services offered by service providers, such as phone, SMS, etc., and instead, use a variety of data services. The service provider must be aware of the user's QoS requirements in this scenario in order to make an informed decision and act appropriately. This requires service providers to make their network available for Over the Top (OTT) services. The provision of OTT through a mobile network presents several difficulties.

Different transmission qualities will be present in certain OTT apps. The heterogeneous Ultra-Dense Network (UDN) features of 5G backhaul networks are followed by OTT on mobile networks. The best way to manage heterogeneous applications is to build numerous logical networks on top of shared physical infrastructure. As a result, network slicing and SDN will have a big influence on network management in mobile networks since the transmission characteristics for a particular UDN vary from user to user or application to application. It is important to note that mobile traffic is more dependent on the day than the night, and depends more on metropolitan regions than on rural ones. Mobile traffic is now non-ergodic due to all of these factors. This has a major and wide-ranging effect.

2. LITERATURE REVIEW

Vineet Raina and Srinath Krishnamurthy (2022) quite frequently, have a natural intuition that tells us whether or not a certain observation is aberrant or unexpected. Sensors that seem to be behaving erratically or giving readings that are rarely seen; an unheard-of combination of symptoms or test readings; a rare pattern in a medical image such as a CT scan; and network traffic in an IT system that is unusual are a few examples of situations that tend to draw attention to themselves. Anomaly detection refers to the process of searching through a set of data in search of any unusual patterns or occurrences.

Vunnava Dinesh Babu and K. Malathi (2022) Deep learning (DL), as it is known today, has become the pattern for machine learning (ML). When compared to NL, DL is the most effective, both in terms of time and money. In DL, there are no restrictions imposed on the various instructional methods. DL algorithms are capable of extracting features of

high quality based on the dataset that is supplied. Because of the rapid advancement of internet technology, enormous amounts of data have recently begun to be processed by being divided into sections according to various criteria. The processing of enormous and complicated datasets, such as pattern matching, recognition of handwriting and voice, analysis of financial markets, and a great many other applications, is more amenable to DL algorithms because of this compatibility. DL has made more progress on a variety of applications, and this will allow it to resolve several challenges that arise in the context of intricate pattern applications. In the past, datasets such as handwritten were utilized in order to locate an accurate result using the Ensemble novel classifier (ENC). However, this did not function to the expected standard. This study presents the design and development of a dynamic deep learning algorithm (DDLA) for the purpose of processing intricate and complex datasets. The Auto Encoder (AE) and the Adaptive Convolutional Neural Network have been combined to create this (A-CNN). The effectiveness of the EDDLA, when combined with ENC, is demonstrated by the experimental findings.

Shrirang Ambaji Kulkarni et al. (2022) it has been demonstrated that ensemble learning algorithms are among the most effective machine learning algorithms in terms of achieving optimal performance in regression and classification problems for several applications. Extreme Gradient Boosting, often known as XGBoost, has become an increasingly popular ensemble learning approach due to the virtues of its performance when applied to structured datasets that are either small or medium in size. In recent years, a potential ensemble technique known as Light Gradient Boosting Machine (Light GBM) has evolved. In terms of performance, this strategy is in direct competition with XGBoost. In addition, Lasso Regression has demonstrated

capability in feature selection and application for use with limited datasets. This paper presents an illustration of an experiment that was carried out on a diabetes dataset. The authors tested the hypothesis that features selection has relatively little of an impact on the performances of Deep Learning Algorithms because these algorithms have the capability, already built-in, of performing feature selection on their own in terms of the layers they contain. Consequently, in order to verify the stacking hypothesis, Deep Learning Multi-Layer Perceptron (DMLP) was used in conjunction with several optimum algorithms such as XGBoost, Light GBM, and Lasso Regression. In the study under consideration, DMLP with all feature variables (DMLP-ALL) achieved a higher R2 than DMLP with stacked chosen features (DMLP-MS) did. This advantage was 8.78%. In addition, DMLP-ALL performed 10.25% better in terms of R2 than the method that was used as a benchmark, Automated Machine Learning (AML). The validation of the proposed stacking models through the use of a dataset of a moderate size offers findings that are encouraging for deep learning models that have been stacked with a powerful Level-0 learner.

Patrick Schneider and Fatos Xhafa (2022) Finding unexpected or unusual occurrences in data streams, often known as anomalous events, is the goal of the field of research known as anomaly detection. The discovery of anomalies may be valuable either on its own or as a starting point for the discovery of additional information hidden in the data. Anomaly detection is an essential component in a wide variety of applications, and it is particularly important in real-time applications. These are the kinds of applications that must be able to identify anomalies in order to function properly; some examples include those dealing with health, critical infrastructure, and security. This chapter provides an overview of the difficulties associated with anomaly

identification and provides examples of the many contexts and features that may be used to categorize abnormalities. This chapter devotes a significant portion of its content to classifying the many approaches of anomaly detection that are now in use so that particular application features may be assessed.

Hai Tao et al. (2022) The recent application of software-defined networking (SDN) in the automotive industry has sparked a lot of interest due to SDN's ability to ensure focused organization across the board in addition to its vast versatility of isolating data from the control plane. This has caused SDN to become a hot topic in recent times. On the other hand, the most significant cause for concern in the traffic network is the fact that numerous emergency and detour notifications are anticipated to be communicated by or brought into linked cars. As a consequence, the revelation of this information leads to an analysis of the direction and control exercised over such messages. Utilizing an SDN-assisted approach in heterogeneous vehicular organizations to execute information traffic using optimum path determination and guidance in scenarios such as a blockage in Vehicle Ad Hoc Networks is the goal of this research (VANETs). In order to make this a reality, a model that is capable of reproducing traffic has been presented, and further tests have resulted in re-enactment findings that better the traditional information flow structure in terms of Round-Trip Time (RTT) and Packet Delivery Ratio (PDR).

3. METHODOLOGY

There are three tiers of service, or "applications," according to IEEE 802.1Q. These are Expedited Forwarding (EF), Assured Forwarding (AF), and Best Effort (BE). EF has the lowest loss, latency, and jitter and is not queued since it is the highest priority. Planning a Quality-of-Service policy allows you to set priorities

for network traffic and services. The prioritized data can subsequently be assigned to a certain AF tier. Apps that use less bandwidth and are less sensitive, such as MULTIMEDIA, will be classified as AF, whereas Emails and P2P applications go under BE. Under EF, we have also seen two distinct kinds of videos: interactive and buffered. Data packets for interactive video are typically between 400 and 800 bytes in size, whereas those for buffered video are often close to the Maximum Transferable Unit's maximum size (MTU). The size of the sound file is around 150-300 bytes. Since MTU size is used for the transport of messages and images, we must rely on additional indicators, such as variation between packets, to determine the nature of the traffic. As traffic patterns shift, however, the interactive video adapts by increasing or decreasing packet size as necessary

4. RESULT AND ANALYSIS

We conducted experiments to evaluate the effectiveness of the suggested approach, MCOP, in terms of throughput and reaction time. Three state-of-the-art approaches (ECMP, HiQoS, and WCMP) are used to evaluate the

suggested method. All of the elephant flow packets are collected by wire shark so that latency, usage, and throughput can be determined. Our experimental setup for this study using the Open Flow network is depicted. We have implemented an SDN using Open Switches and a Floodlight controller. The connection has been established, and there will be 30 Mbps of level 1 bandwidth and 10 Mbps of level 2 bandwidth hosts connect to switches at level 1, and backbone connections are used at level 2. From node 1 through node 7, we create two elephant flows, one using EF and the other using AF. Bandwidth requirements are significant for an elephant flow. To show BE traffic, a small number of flows, between 0.5 and 5 Mbps, are produced at random and forwarded by other nodes. Whenever a packet enters the network, it first checks the table to see if any matching rules exist, and if not, it is sent to the controller so that the controller may make any required decisions and set up any necessary policies. To begin with, we determined the maximum and minimum latency between the pathways and determined their deference. The value we obtained was 80 ms. at regular 3-second intervals; our RM component records the network connection. At node 1, we use the pktgen tool to transmit two different kinds of applications. We have evaluated MCOP's efficacy for latency- and bandwidth-sensitive applications including voice and video (AF) (EF).

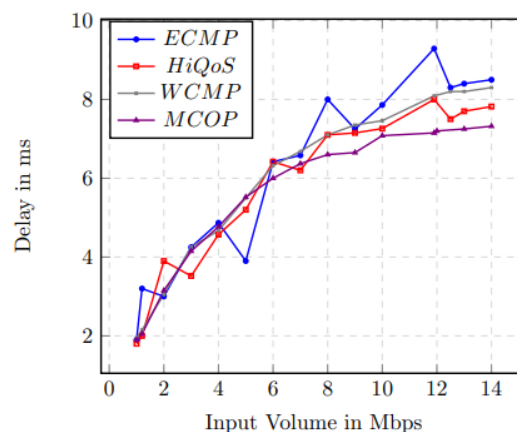


Figure 1: Delay Comparison between MCOP with WCMP, ECMP, and HiQoS for traffic type EF

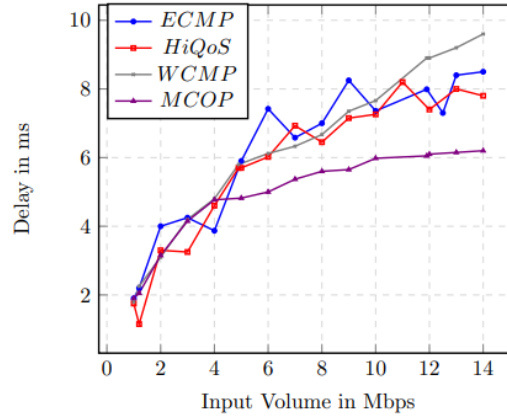


Figure 2: Delay Comparison between MCOP with WCMP, ECMP, and HiQoS for Traffic Type AF

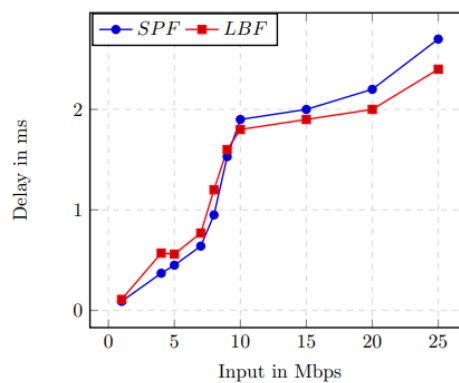


Figure 3: Delay comparison Between SPF and LBF

4.

1 Effective Evaluation

ECMP, HiQoS, and WCMP are three state-of-the-art approaches that we have compared our work with for delay-sensitive and bandwidth-sensitive applications, respectively. Node 1 produces at a rate of 5-40Mbps with an Ethernet frame size of 1000 bytes, making it suitable for bandwidth-sensitive applications (AF) and delay-sensitive applications (EF).

- **Delay:** Delay comparisons between

our technique and HiQoS, WCMP, and ECMP for AF and EF applications are displayed in Figures 1 and 2. At low loads, the available bandwidth is adequate to send all data without any need to divide the traffic. Our method achieves the same performance under low and high loads by triggering multipath when input volume grows and network congestion occurs. Our method outperforms ECMP and HiQoS for bandwidth-hungry AF applications.

This is because, unlike HiQoS and WCMP, our method uses a flow allow where TCP packet reordering is unnecessary. After the connection becomes congested, ECMP automatically sets up several pathways. However, the appropriate forwarding path must be determined for a bandwidth-intensive application. In our method, this is a criterion. In a similar vein, Figure 1 displays that the output from HiQoS, WMCP, and MCOP is almost identical when applied to the voice application. Our method MCOP, however, produces somewhat

superior results. Our method uses SPF, the quickest route for transmitting EF with QoS. The delayed performance of SPF and LBF has also been evaluated. In Figure 3, we can see how SPF and LBF do in terms of delay performance when the switches are under a heavy load, that is, when there are many more policies in the switches. At a reduced data rate, SPF performs well for the application. A better reaction from LBF may be seen when the input traffic rate rises. Greater numbers of plans anticipate greater numbers of loads.

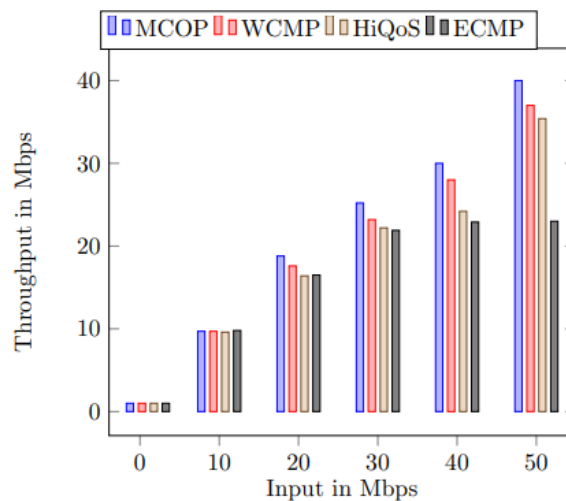


Figure 4: Throughput Performance Comparison for Traffic type AF with MCOP, WCMP, ECMP, and HiQoS

- Throughput:** Throughput at the target system for various approaches is measured in this experiment. Good throughput was obtained at the end system for both AF and EF, as shown in Figures 4 and 6. Until the first chosen backbone link becomes saturated, all approaches yield the same outcome; after that, our approach and the HiQoS begin multipathing. Due to the fact that ECMP across many pathways is capped at the lowest

available bandwidth, its throughput is lower. For both AF and EF traffic, our MCOP technique provides a more immediate reaction. Our method maximizes the utilization of available bandwidth to prevent the load imbalance and its negative effects on latency and throughput. Since our method employs SPF or LBF bucket if available after determining the kind of application, as indicated, we achieve a significant improvement in

performance for BE traffic. Comparing SPF vs LBF in terms of throughput is something we have also investigated. Once multipath selection begins for the input load, the output improves in LBF (see Figure 5). Computer simulation is an abstract representation of a real-world system that can be made to behave like the actual thing by applying the model's specifications. In order to keep the model manageable, it is necessary to focus on only a subset of the system's traits, attributes, or behaviors when doing a simulation. The following criteria are taken into account. The bandwidth is 30 Mbps at

level 1 and 10 Mbps at level 2 with a latency of 0.33 ms and 1 ms, respectively. The buffer size at the node is 1 MB, whereas at the intermediate switches it is 64 KB. The consequent boost in TCP-based AF applications is substantial. This is due to the fact that MCOP reorders packets. However, ideas like UDP discard the old packets, making reordering unnecessary in other applications like EF. Performance-wise, MCOP is on par with other state-of-the-art approaches but somewhat better than the previous two.

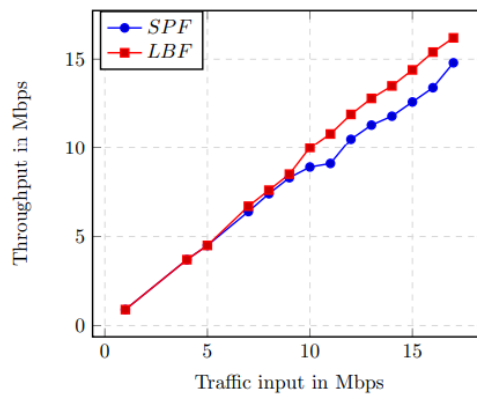


Figure 5: Throughput Comparison between LBF and SPF

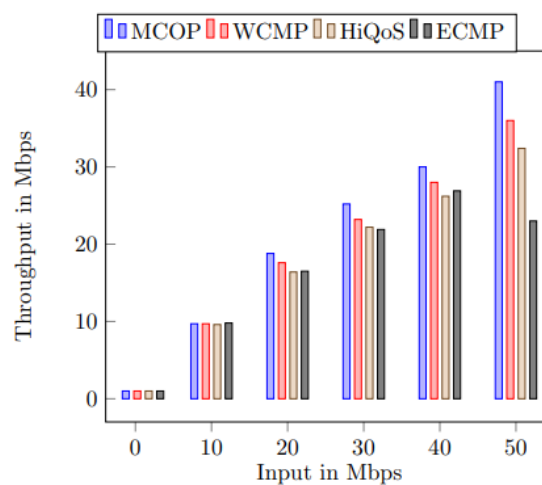


Figure 6: Throughput performance comparison for traffic type EF with MCOP, WCMP,

ECMP, and HiQoS

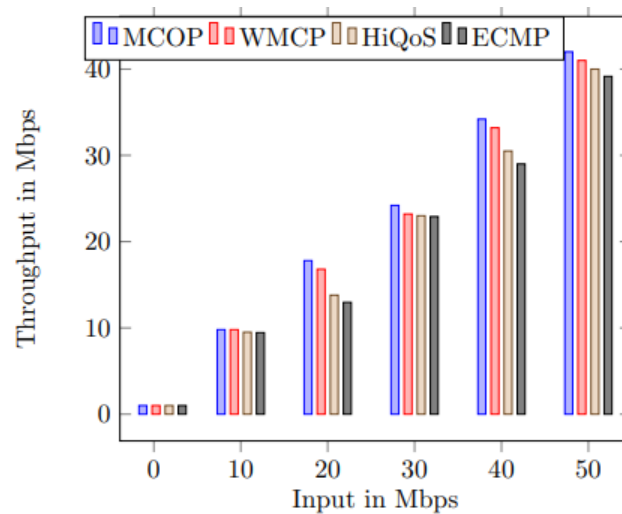


Figure 7: Throughput Performance Comparison for Traffic type BE with MCOP, WCMP, ECMP and HiQoS

In addition, a threshold is used in both HiQoS and WCMP to pick a path based on the "price" components. If the calculated loss is less than the threshold, standard forwarding is used. Maximum throughput (TF) between a source and a destination is used to choose which path to use in MCOP. Since the situation under consideration has more pathways while still satisfying the requirements for MCOP and HiQoS, the outcome is not drastically different. In any case, the MCOP anticipates the path at the outset using the TI model and allocates for the request, so that no extra messages need to be sent between the networks.

5. CONCLUSION

To accomplish multipath without identifying a source, suitable rules may be deployed at the switches using flow allow. Also determined is the present load on the specified route. Throughput increases because better load balancing is made possible by this. Therefore, applications that are particularly sensitive to delays, such as those in the conversational and streaming

classes, are typically associated with real-time traffic, such as video-telephony, telnet, audio, and video services. Bandwidth load balancing is a useful tool for optimizing network performance.

REFERENCES

1. Raina, Vineet & Krishnamurthy, Srinath. (2022), "Anomaly Detection," 10.1007/978-1-4842-7419-4_8
2. V. D. Babu and K. Malathi (2022), "Dynamic Deep Learning Algorithm (DDLA) for Processing of Complex and Large Datasets," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), pp. 336-342, doi: 10.1109/ICAIS53314.2022.9743013
3. S. A. Kulkarni, V. P. Gurupur and C. King (2022), "Impact Analysis of Stacked Machine Learning Algorithms Based Feature Selections for Deep Learning Algorithm Applied to Regression Analysis," SoutheastCon

- 2022, pp. 269-275, doi:
10.1109/SoutheastCon48659.2022.976
4105
4. Schneider, Patrick & Xhafa, Fatos. (2022),” Anomaly detection,” 10.1016/B978-0-12-823818-9.00013-4
 5. Hai Tao et al. (2022),” SDN-assisted technique for traffic control and information execution in vehicular ad-hoc networks,” Computers and Electrical Engineering, Volume 102, 108108
 6. Kim, Haesik. (2022),” Deep Learning,” 10.1007/978-3-030-95041-5_6
 7. Setoodeh, Peyman & Habibi, Saeid & Haykin, Simon. (2022),” Deep Learning,” 10.1002/9781119078166.ch8
 8. Sunday, Uzoma. (2021),” Software Defined Networking,” Telecommunications and Radio Engineering
 9. Alzubaidi, L., Zhang, J., Humaidi, A.J. et al.(2021),” Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” J Big Data 8, 53, <https://doi.org/10.1186/s40537-021-00444-8>
 10. Sarker, I.H. (2021),” Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions,” SN COMPUT. SCI, 2, 420, <https://doi.org/10.1007/s42979-021-00815-1>

Author’s Declaration

I as an author of the above research paper/article, hereby, declare that the content of this paper is prepared by me and if any person having copyright issue or patent or anything otherwise related to the content, I shall always be legally responsible for any issue. For the reason of invisibility of my research paper on the website/amendments /updates, I have resubmitted my paper for publication on the same date. If any data or information given by me is not correct I shall always be legally responsible. With my whole responsibility legally and formally I have intimated the publisher (Publisher) that my paper has been checked by my guide (if any) or expert to make it sure that paper is technically right and there is no unaccepted plagiarism and the entire content is genuinely mine. If any issue arise related to Plagiarism / Guide Name / Educational Qualification / Designation/Address of my university/college/institution/ Structure or Formatting/ Resubmission / Submission / Copyright / Patent/ Submission for any higher degree or Job/ Primary Data/ Secondary Data Issues, I will be solely/entirely responsible for any legal issues. I have been informed that the most of the data from the website is invisible or shuffled or vanished from the data base due to some technical fault or hacking and therefore the process of resubmission is there for the scholars/students who finds trouble in getting their paper on the website. At the time of resubmission of my paper I take all the legal and formal responsibilities, If I hide or do not submit the copy of my original documents (Aadhar/Driving License/Any Identity Proof and Address Proof and Photo) in spite of demand from the publisher then my paper may be rejected or removed from the website anytime and may not be consider for verification. I accept the fact that as the content of this paper and the resubmission legal responsibilities and reasons are only mine then

the Publisher (Airo International Journal/Airo National Research Journal) is never responsible. I also declare that if publisher finds any complication or error or anything hidden or implemented otherwise, my paper may be removed from the website or the watermark of remark/actuality may be mentioned on my paper. Even if anything is found illegal publisher may also take legal action against me.

Rounak R. Gupta
Dr. Pramod Sharma
