IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Implementation of Reconfigurable and Dynamic QoS Provisioned Optical DCN Architecture for Mixed Traffic Environment

[1] Sougata Bera, [2] Chandi Pani, [3] Sayantani Sen

[1] Assistant professor, [2] Professor, [3] Assistant professor
[1]Department of Computer Science and Engineering,
[1] Meghnad Saha Institute of Technology, Kolkata, India

Abstract: This paper presents a novel heuristic algorithm designed to optimize path reconfiguration and dynamically provision Quality of Service (QoS) within Passive Optical Device Switch (PODS)-based Data Center Networks (DCNs). In response to the increasing demand for high-bandwidth and low-latency services, particularly in environments characterized by mixed traffic types, the proposed solution aims to overcome the limitations associated with traditional packet-switching networks. The PODS architecture, which integrates an Arrayed Wave Guide Router (AWGR) and a control unit, intelligently manages the allocation of packets into service-specific buffers. The heuristic algorithm enhances this process by reallocating packets to alternate buffers within the same or different service classes when the initially requested buffer is unavailable. Furthermore, it dynamically reconfigures the transmission path to ensure uninterrupted service, thereby maintaining QoS standards across varying traffic loads.

To validate the efficacy of this approach, a laboratory test bench model was constructed using 7 Top of Rack (ToR) switches connected to Raspberry Pi modules, simulating a realistic data center environment. The experimental results reveal that the optimal configuration is achieved with 14 buffers per ToR switch, resulting in 0% blocking for high-priority real-time traffic and an overall blocking probability of just 18.3%. These findings underscore the potential of the PODS-based architecture to significantly enhance network performance, scalability, and reliability in next-generation data centers.

***Index Terms* -** Optical DCN, AWGR, QoS Provisioning, Wavelength Assignment

I. Introduction

Next-generation Data Centers are expected to adopt new architectures and technologies to enhance their performance in terms of throughput, latency, scalability, and power consumption [1]. Optical data center networks rely on various optical switching technologies, including SOA-based switches, Micro-Electro-Mechanical Systems (MEMS) switches, and Arrayed Waveguide Grating Routers (AWGR). MEMS switches, used in systems like c-through [2] and Helios [3], are reconfigurable optical switches driven by power but have a long reconfiguration time, making them less suitable for fast packet switching in data center networks. Several hybrid electro/optical interconnecting architectures have been proposed for data center networks [4]. Most of the architecture required a centralized scheduler to reconfigure the entire architecture in response to traffic dynamics except RotorNet [5]. It has a predefined scheduler, but it is not sensitive to traffic dynamics. AWGR is an optical device that resolves packet contention in the wavelength domain through its cyclic routing characteristic. It allows multiple inputs to reach the same output simultaneously. Several AWGR-based data center network architectures, such as DOS [6] and Petabit [7], use tunable wavelength converters (TWCs) for flexible wavelength management. However, TWCs are power-hungry and consume significant electrical power during operation. In this domain, contention resolution is achieved by wavelength conversion, fiber optic delay lines (FDLs) and rarely used deflection routing as mentioned in reference [8]. The Passive

Optical Data Center Network Architecture (PODCA) [9] is a passive optical DCN architecture developed using AWGR and TWC, controlled by a Control Unit that assigns desired wavelengths. It calculates wavelength without level extraction of packets and has a packet latency below 9 μs compared to other passive optical DCN architectures like DOS and LIONS [10]. Recently proposed and investigated architecture OPSquare [11,12]and ROTOS [13], an optical DCN architecture, utilizing fast optical switches and optical flow control, offering low latency, high connectivity, low cost, and power consumption.

Flexible quality of service provisioning and scalable scheduling are the other challenges for network design engineers. Several solutions are proposed that can scale the network and address the system challenges such as demand estimation and bandwidth reconfiguration [14]. For fast and scalable switch control a synergistic switched control optical network [15] is proposed that processes the label signal and configures path within the nanosecond range. An SDN-enabled solution [16] addressed contention resolution problems in the optical domain by flow control.

In this paper, a test bench DCN architecture is implemented that integrates fast and scalable switching, and dynamic QoS provisioning features in a single architecture module. The proposed heuristic algorithm is run over the proposed model to support path reconfiguration and dynamic QoS. According to algorithm, the packets from the servers are stored in a priority buffer, if the buffer of the requested service class is not available then it will move to the next available buffer before forwarding, and a loopback method is used to reroute the packet if the wavelength is not available for forwarding the packet to the proper destination. The packet transmission priority is set by dynamic assignment, meaning once the higher priority packet arrives it will directly forward to the proper destination without retransmission or loopback methodology.

The rest of the paper is organized as follows: Section II describes the system model and its working principle. The test bench experimental model of PODS-based DCN architecture using Raspberry Pi is described in section III. The result analysis is presented in section IV. And finally, section V is the conclusion of the paper.

II. System Model and Working Principle:

The proposed Passive Optical Datacenter Switch (PODS) base DCN architecture with a control unit [17] is shown in Fig.1. The proposed model consists of TOR, AWGR, TX and RX modules. Each TOR is connected to the end users or servers. The rest of the ports of the TOR are configured as input ports and output ports respectively. The input port of the AWGR is connected to the TOR via TX module and the output port of the AWGR is connected to the TOR via the RX module respectively. The TX module of each ToR consists of an electrical buffer (EB), optical channel adapter, optical label generator (OLG), packet encapsulator (PE), and electro-optic converter (lasers) to send the incoming packets to the AWGR through tunable wavelength converter (TWC). After passing through the TWC, wavelengths are combined by an Optical Multiplexer (OMUX) and finally reach to the input port of the AWGR. The RX module consists of an Optical Demultiplexer (ODMUX) followed by an optical receiver called an optical-to-electrical converter (OE converter), electrical buffer (EB), and packet adapter (PA).

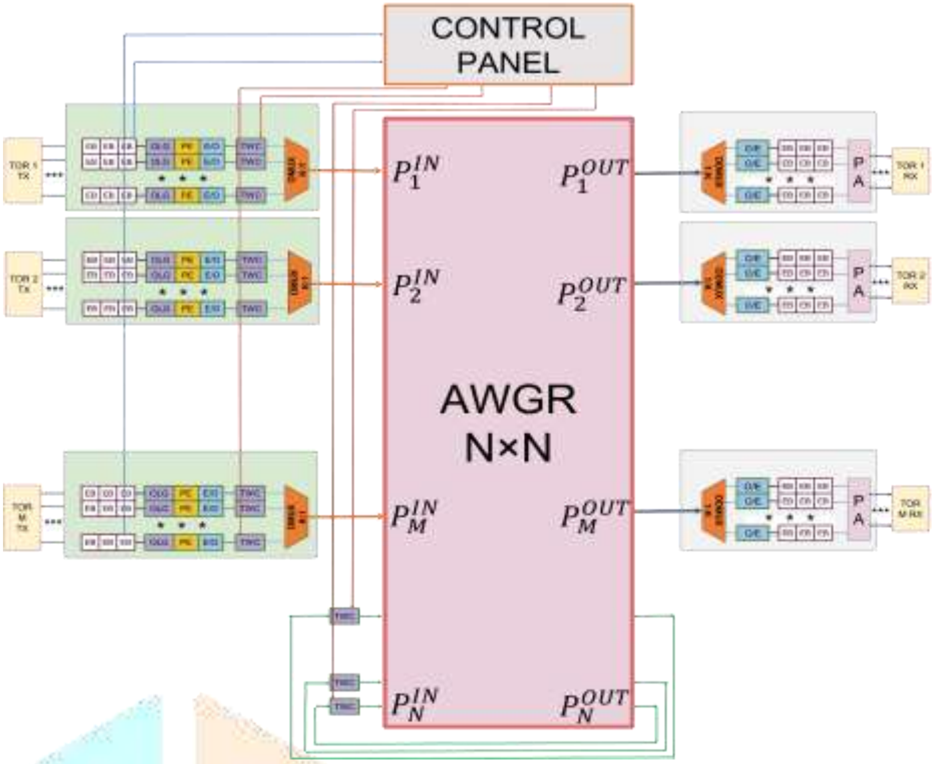


Fig. 1 PODS-Based DCN Architecture

In this architecture, the generated packets from the server, first arrive to any input TOR (ToR_{IN}), then as per the service class of the packet, they are placed in the shared buffer marked as EB. Similarly, when the packets are out from the ToR_{OUT} port of AWGR, after demultiplexing the packets they are converted from optical to electrical and stored in a shared buffer of TOR receiver (RX) module. As per the service class of the packet, the priority level of the buffer is set. In this paper we only consider four types of service class traffic, and based on it buffers are classified as high-priority real-time (HRT), standard-priority real-time (SRT), Earliest Deadline First (EDEL), First-Come-First-Served (FCFS). Round-robin processing is used to handle packets from the buffer. The number of buffers allocated to any service class is based on the incoming traffic requirement. Thus, the optimal architecture is determined by allocating a certain number of buffers under each service class in accordance with the traffic pattern. The architecture's performance is then evaluated in terms of blocking probability by reusing the buffer of a different service class (if it is empty) and rerouting wavelengths using the loopback method. Fig. 2 shows the flow chart for the proposed algorithm and the function of the PODS-based DCN architecture.

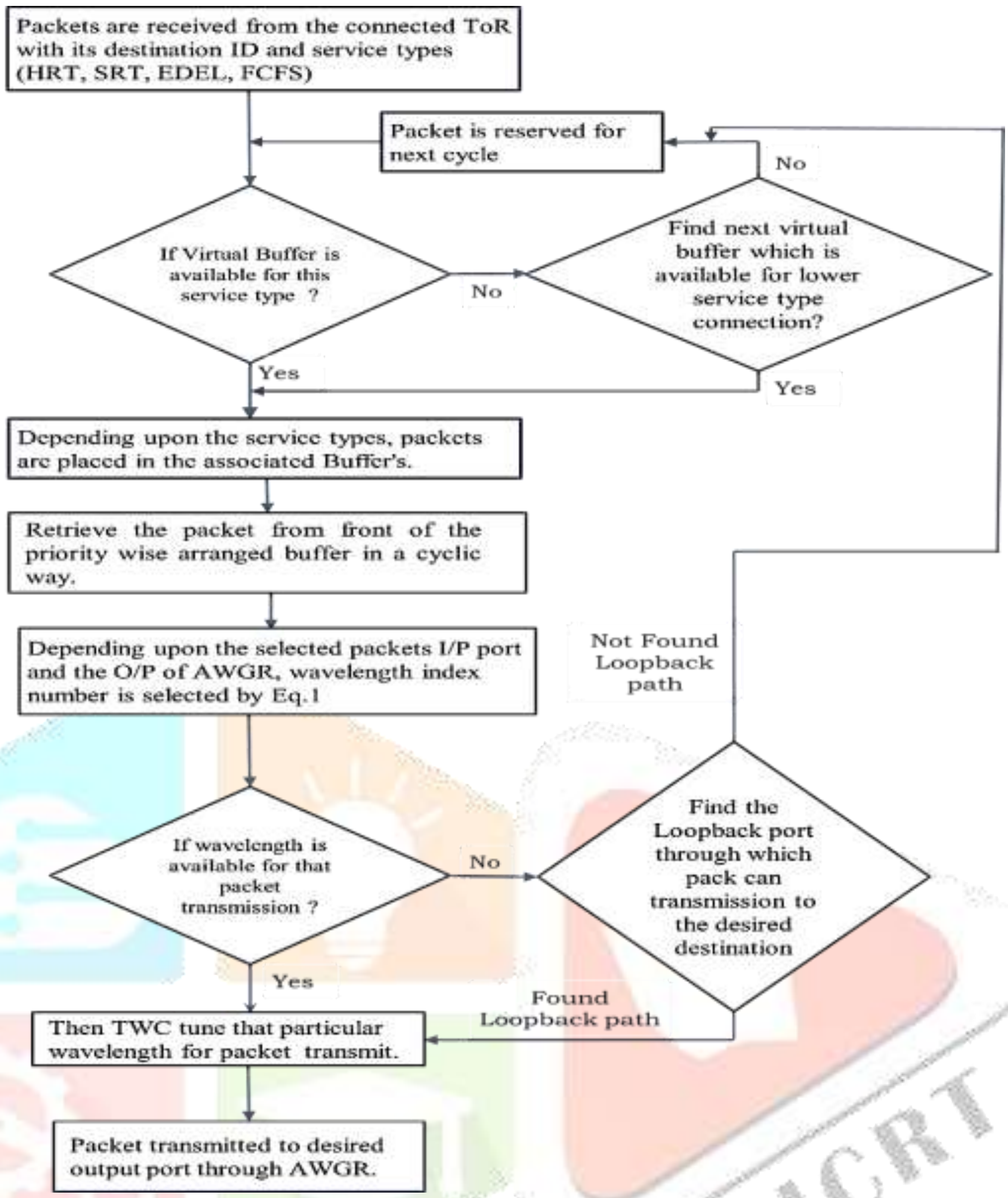


Fig. 2 Flowchart of PODS-Based DCN Architecture

III. Test Bench Implementation of the proposed model using Raspberry Pi:

To find the performance of the proposed architecture and algorithm in the hardware platform, a test bench is developed in our laboratory with 7 servers. The AWGR and the control unit are implemented in Raspberry Pi. We use Broadcom BCM2711, Quad-core Cortex-A72 (ARM v8) 64-bit Raspberry Pi module to implement the 8x8 AWGR and Control Unit. Out of 40 GPIO ports of the Raspberry Pi module, only 7 ports are connected to 7 Servers (Configuration: Intel Core i7 - 12700T, 1.40 GHz, 8GB DDR4 3200MHz RAM, SSD - 512 GB) that are used as TOR. Two ports are used as a loopback path for input and output. Another 7 ports are connected to the output port of the TOR. Two ports are used as a loopback path for input and output. Another 7 ports are connected to the output port of the TOR. The detailed pin configuration for Input and output ports are as follows: PIN 11, 13, 15, 29, 31, 35, and 37 act as Input ports of AWGR, and PIN 16, 18, 22, 32, 36, 38, 40 acts as the output port of AWGR respectively, which are connected to the TOR (Server). The block diagram of the test bench setup is shown in Fig. 3. Separate memory of 7GB (extended up to 14 GB) is allocated within the Raspberry Pi to effectively buffer the incoming packets. This memory is divided into each of the assigned input ports, providing 1GB of buffered memory for each port. This 1GB is further subdivided into B1, B2, B3 and B4 of variable size. Fig. 4 shows the buffer allocation scheme in the proposed test bench implementation. Fig. 5 shows the Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit Raspberry-Pi module and its pin configurations and Fig.6 shows the test bench implementation model of the proposed architecture.

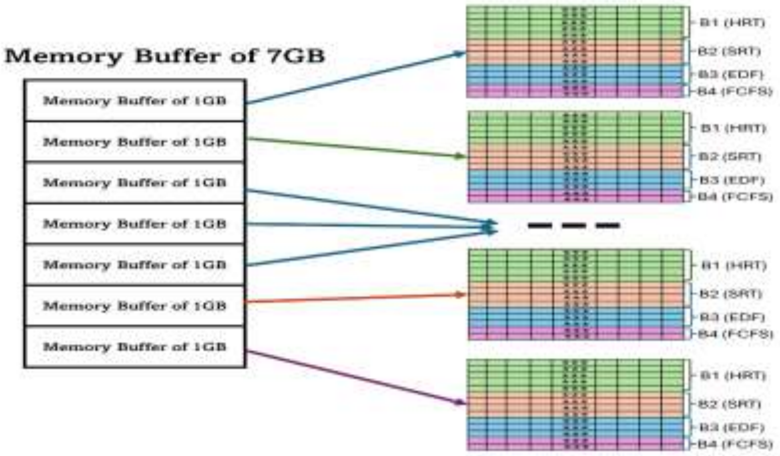


Fig.3 Block Diagram of Test Bench setup

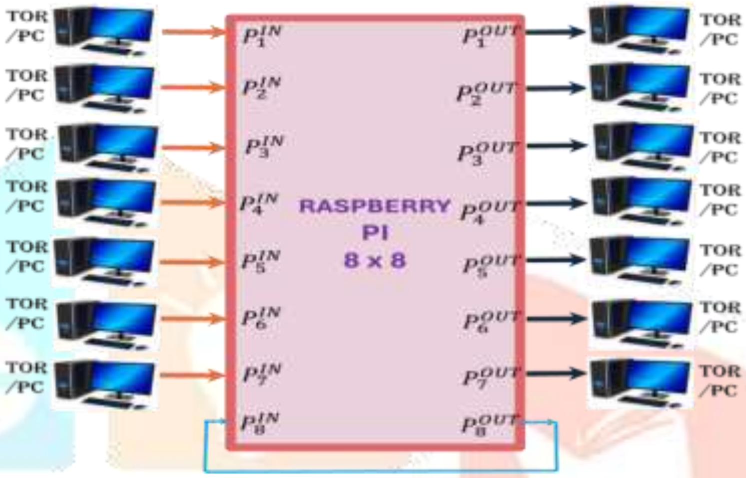


Fig. 4 Buffer allocation scheme in the proposed test bench implementation

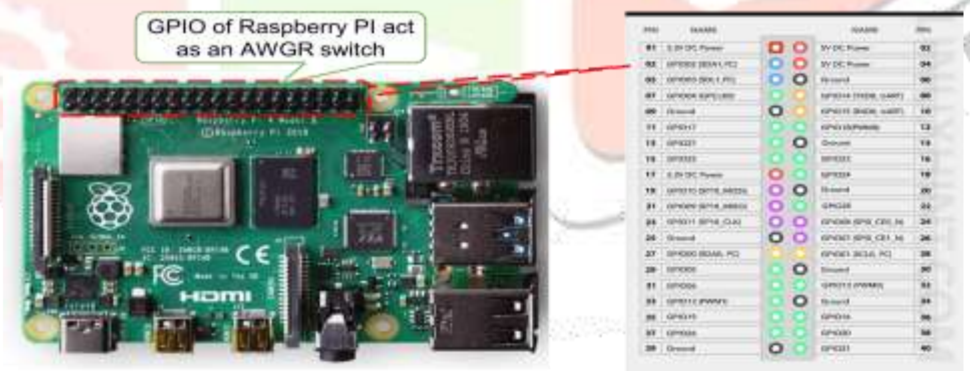
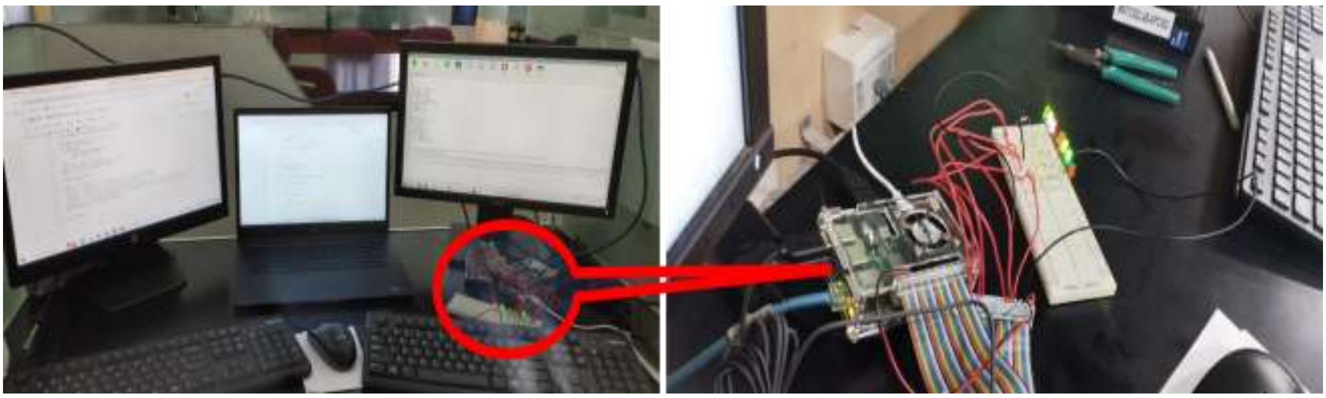


Fig. 5 Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit Raspberry-Pi module and its pin



configurations

Fig. 6 shows the test bench implementation model of the proposed architecture

IV. Performance Analysis and Results:

To Study the performance of the network we use the Integer Linear Programming (ILP) model to optimize the allocation of buffers and wavelengths for packet transmissions within a time slot. The algorithm is developed in the Python platform and integrated with the Raspberry Pi module for 16 wavelengths. ToR-to-ToR communication is established through the Raspberry Pi module via a socket connection. In the experimental setup, each ToR can transfer multiple files parallelly to different locations. So, in the implementation model, a total of 112 (7x16) concurrent connections are possible. For performance analysis network load is increased from 50% (56 connections) to 100% (112 connections) and the blocking probability of the entire network is calculated by varying the number of buffers under each ToR, as well as placing the packet in underutilized buffers of other service classes. For the experimental purpose initially, we started with 8 no. of buffers per ToR. This buffer per ToR and its subdivisions under each service type is mentioned in Table 1. For simulation purposes, we divide the network load with the service type B1=35%, B2= 30%, B3=20% and B4=15%. Then the number of buffers is dynamically changed from 8 to 16 and packet allocations are also changed keeping the network load remain the same. In analysis, first, the blocking probability is calculated without buffer reuse and with buffer reuse. For each measurement, 100 time stamps are taken, and the results are computed by averaging all the values.

Table 1. Buffer Allocation based on service type in terms of network Load per ToR

NO OF BUFFER	(B1) 35%	(B2) 30%	(B3) 20%	(B4) 15%
8	3	2	2	1
9	3	3	2	1
10	3	3	2	1
11	4	3	2	2
12	4	3	3	2
13	4	4	3	2
14	5	4	3	2
15	5	4	3	2
16	5	5	3	2

Fig 7 shows the snapshot of the execution of the program and wavelength assignment for all ToRs. In Fig. 7 marked at block A indicates that there are 98 connections are generated in this instant and out of it 80 connections are successfully established, and 18 connections are not possible due to a lack of buffer availability and wavelength allocation conflict. In Fig.7 marked as block B indicates the allocation of wavelengths. The sequence [1,1,1,1,1,1,1,1,0,1,0,1,1,0,1,1] delineates the allocation of wavelengths in port 7. The wavelengths w_1 to w_8, w_{10}, w_{12}, w_{13} , w_{15} and w_{16} are actively utilized means connections are properly established using these wavelengths, while wavelengths w_9, w_{11}, and w_{14} remain unutilized. Block C of Fig. 7 indicates the value of wavelengths that are utilized for successful communication

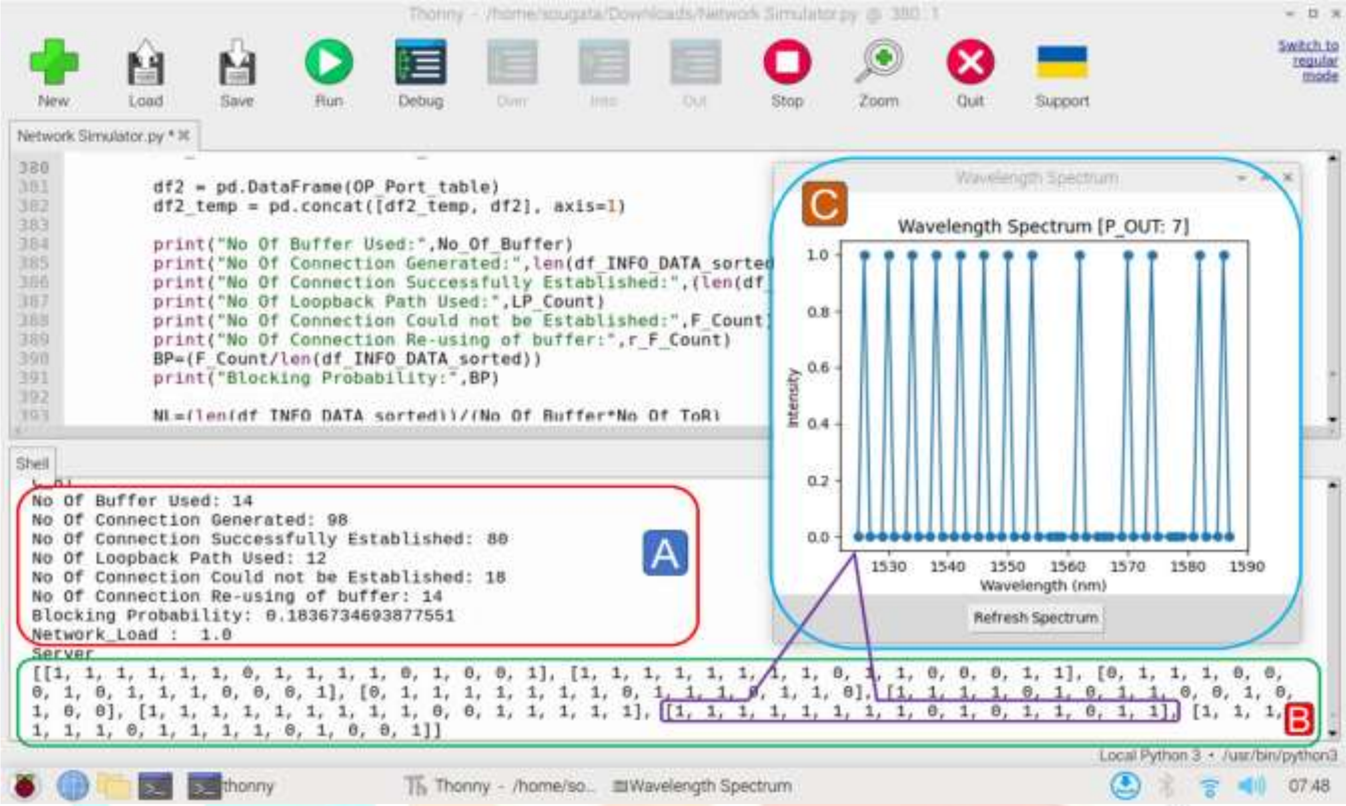
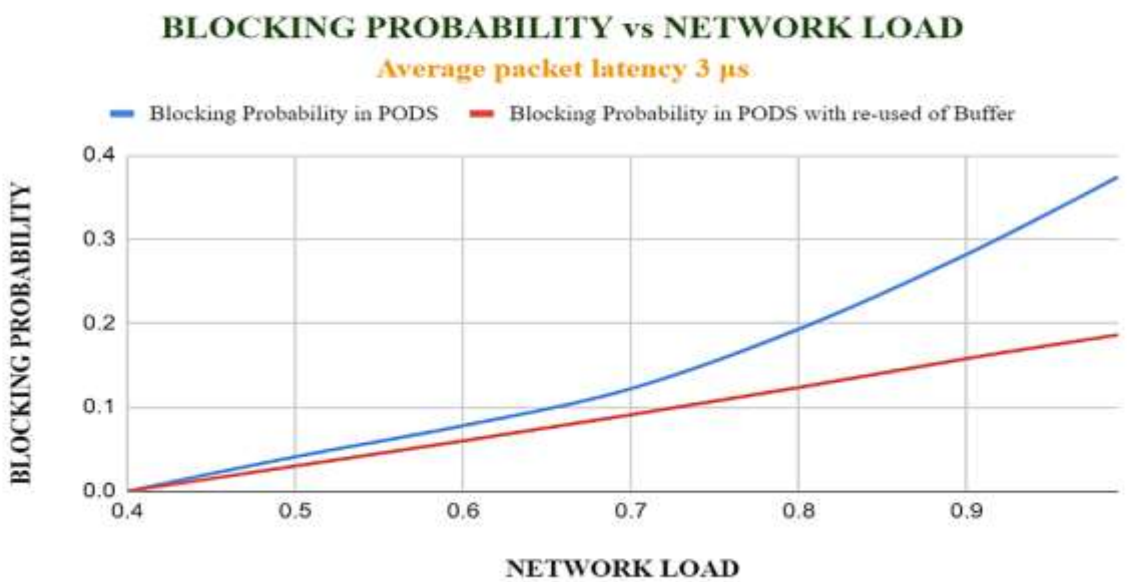


Fig. 7 Snapshot of execution of program and wavelength assignment for the architecture

Table 2 shows the blocking probability of the network without buffer reuse and with buffer reuse. Fig. 8 shows the blocking probability with network load, with and without reused buffers, and found an 18.3% improvement in blocking probably in the case of the reused buffer. Table 3 shows the blocking probability (with reused buffer) with buffer size for different network loads. Fig. 9 shows the blocking probability with respect to no of buffer for different network loads. From Fig. 8 it is observed that blocking probability remains constant for the buffer size 14 onwards. So, for further calculation, we optimized our network with 14 buffers per port.

Table 2. Blocking probability with respect to network load with different buffer size

NETWORK LOAD	0.4	0.5	0.6	0.7	0.8	0.9	0.99
Blocking Probability in PODS without re-used buffer	0	0.041	0.078	0.122	0.193	0.282	0.374



Blocking Probability in PODS with re-used of buffer	0	0.03	0.06	0.091	0.1235	0.158	0.186

Fig. 8 Blocking probability with respect to network load

Table 3. Blocking probability with respect to network load with different buffer size

NETWORK LOAD		**50%**	**56%**	**63%**	**69%**	**75%**	**81%**	**88%**	**94%**	**100%**
BUFF ER SIZE	**8**	0.170	0.206	0.243	0.286	0.333	0.385	0.418	0.452	0.482
	9	0.107	0.138	0.171	0.221	0.262	0.316	0.352	0.390	0.420
	10	0.068	0.087	0.114	0.162	0.202	0.253	0.294	0.326	0.357
	11	0.043	0.060	0.086	0.125	0.161	0.201	0.244	0.271	0.295
	12	0.033	0.048	0.071	0.104	0.131	0.165	0.200	0.224	0.246
	13	0.030	0.048	0.071	0.091	0.119	0.143	0.171	0.189	0.207
	14	0.030	0.048	0.071	0.091	0.113	0.132	0.153	0.171	0.188
	15	0.030	0.048	0.071	0.091	0.113	0.126	0.143	0.162	0.179
	16	0.030	0.048	0.071	0.091	0.113	0.126	0.143	0.162	0.179

From Fig.7 we see that 18 connections are not possible to establish this specific instance. The distribution of blocking under various service classes is displayed in Table 4. From the table, it is observed that no blocking is found for HRT traffic, 16% of total blocking in SRT, 32% of total blocking in EDF, and 50% blocking in FCS, and overall blocking is 18.3%.

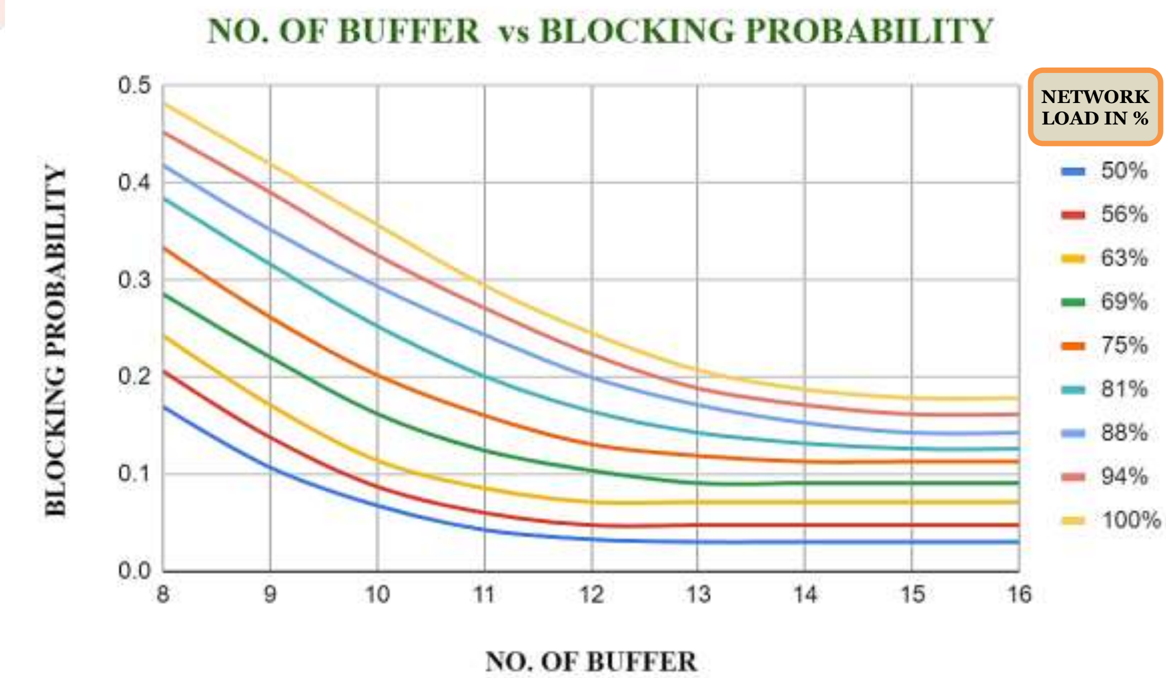


Fig. 9 Blocking probability with buffer size for different network loads

Table 4. Number of blocked connections in different services

SERVICE TYPE	B1 (HRT)	B2 (SRT)	B4 (EDF)	B4 (FCFS)
Number of block connections with re-used buffer	0	3	6	9

V. Conclusion

The pursuit of fully bisection bandwidth networks through packet switches has shown to increase system complexity without necessarily providing the best performance, especially in environments with mixed traffic types. This study has explored the alternative potential of optical circuit switching (OCS) technology within data center networks (DCNs). By focusing on the critical aspect of blocking probability—particularly for high-priority real-time traffic—this research underscores the ability of OCS to deliver superior quality of service while also simplifying network design.

A heuristic algorithm was developed and integrated into a PODS-based DCN architecture, demonstrating how OCS can effectively manage traffic with minimal complexity. To validate the practicality of our approach, we constructed a prototype model using Raspberry Pi, simulating the real-world application of the proposed solution. The prototype results affirmed the algorithm's capability to significantly reduce blocking probability, making it particularly advantageous in environments with a mix of real-time and non-real-time traffic.

Moreover, the study highlights the broader implications of adopting optical circuit switching in future data center networks. The proposed approach offers a path to achieving greater robustness, scalability, higher data rates, and improved bandwidth efficiency. As data centers continue to evolve, the integration of OCS could be a critical step in addressing the growing demands for speed, reliability, and efficiency in network infrastructures.

References

[1] IEA (2024), Global data center electricity use to double by 2026, IEA, Paris https://www.datacenterdynamics.com/en/news/global-data-center-electricity-use-to-double-by-2026-report/

[2] Guohui Wang, David G. Andersen, Michael Kaminsky, Konstantina Papagiannaki, T.S. Eugene Ng, Michael Kozuch, and Michael Ryan. 2010. C-Through: part-time optics in data centers. SIGCOMM Comput. Commun. Rev. 40, 4 (October 2010), 327–338.

[3] Nathan Farrington, George Porter, Sivasankar Radhakrishnan, Hamid Hajabdolali Bazzaz, Vikram Subramanya, Yeshaiahu Fainman, George Papen, and Amin Vahdat. 2010. Helios: a hybrid electrical/optical switch architecture for modular data centers. SIGCOMM Comput. Commun. Rev. 40, 4 (October 2010), 339–350.

[4] Balanici, M.; Pachnicke, S. Hybrid Electro-Optical Intra-Data Center Networks Tailored for Different Traffic Classes. J. Opt. Commun. Netw. 2018, 10, 889–901, https://doi.org/10.1364/jocn.10.000889.

[5] Mellette, W.M.; McGuinness, R.; Roy, A.; Forencich, A.; Papen, G.; Snoeren, A.C.; Porter, G. Rotornet: A scalable, low-complexity, optical datacenter network. In Proceedings of the Conference of the ACM Special Interest Group on Data Communication, New York, NY, USA, 07 August. 2017, https://doi.org/10.1145/3098822. 3098838.

[6] X. Ye, Y. Yin, S. B. Yoo, P. Mejia, R. Proietti, and V. Akella, "DOS: a scalable optical switch for datacenters," in 6th ACM/IEEE Symposium on Architectures for Networking and Communications Systems, ACM, 2010, p. 24.

[7] K. Xi, Y.-H. Kao, M. Yang, and H. Chao, "A petabit bufferless optical switch for data center networks," in Optical Interconnects for Future Data Center Networks, Springer, 2013, pp. 135–154.

[8] Singh, A.; Tiwari, A.K. Analysis of Hybrid Buffer Based Optical Data Center Switch. J. Opt. Commun. 2018, 42, 415–424, https://doi.org/10.1515/joc-2018-0121

[9] Xu, Maotong & Liu, Chong & Subramaniam, S.. (2018). PODCA: A passive optical data center network architecture. Journal of Optical Communications and Networking. 10. 409. 10.1364/JOCN.10.000409.

[10] Y. Yin, R. Proietti, X. Ye, C. J. Nitta, V. Akella and S. J. B. Yoo, "LIONS: An AWGR-Based Low-Latency Optical Switch for High-Performance Computing and Data Centers," in IEEE Journal of Selected Topics in Quantum Electronics, vol. 19, no. 2, pp. 3600409-3600409, March-April 2013, Art no. 3600409, doi: 10.1109/JSTQE.2012.2209174.

[11] F. Yan, W. Miao, O. Raz and N. Calabretta, "Opsquare: A flat DCN architecture based on flow-controlled optical packet switches," in Journal of Optical Communications and Networking, vol. 9, no. 4, pp. 291-303, April 2017, doi: 10.1364/JOCN.9.000291.

[12] Xue, X.; Wang, F.; Agraz, F.; Pages, A.; Pan, B.; Yan, F.; Guo, X.; Spadaro, S.; Calabretta, N. SDN-Controlled and Orchestrated OPSquare DCN Enabling Automatic Network Slicing With Differentiated QoS Provisioning. J. Light. Technol. 2020, 38, 1103–1112, https://doi.org/10.1109/jlt.2020.2965640.

[13] Xue, X.; Yan, F.; Prifti, K.; Wang, F.; Pan, B.; Guo, X.; Zhang, S.; Calabretta, N. ROTOS: A Reconfigurable and Cost-Effective Architecture for High-Performance Optical Data Center Networks. J. Light. Technol. 2020, 38, 3485–3494, https://doi.org/10.1109/jlt.2020.3002735

[14] Xue, X.; Nakamura, F.; Prifti, K.; Pan, B.; Yan, F.; Wang, F.; Guo, X.; Tsuda, H.; Calabretta, N. SDN enabled flexible optical data center network with dynamic bandwidth allocation based on photonic integrated wavelength selective switch. Opt. Express 2020, 28, 8949–8958, https://doi.org/10.1364/oe.388759.

[15] Xue, X.; Calabretta, N. Synergistic Switch Control Enabled Optical Data Center Networks. IEEE Commun. Mag. 2022, 60, 62–67, https://doi.org/10.1109/mcom.001.2100683.

[16] Xue, X.; Wang, F.; Chen, S.; Yan, F.; Pan, B.; Prifti, K.; Guo, X.; Zhang, S.; Xie, C.; Calabretta, N. Experimental Assessments of SDN-Enabled Optical Polling Flow Control for Contention Resolution in Optical DCNs. J. Light. Technol. 2020, 39, 2652–2660, https://doi.org/10.1109/jlt.2020.3042820

[17] Sougata Bera, Chandi Pani, "Re-routable and Low-latency All Optical Switching Algorithm for Next Generation DCN" - IJFMR Volume 5, Issue 6, November-December 2023. DOI 10.36948/ijfmr.2023.v05i06.9965