

Kalman Filtering-based Traffic Prediction for Software Defined Intra-data Center Networks

Jacques Mbous^{1*}, Tao Jiang¹, Ming Tang^{1,2}, Songnian Fu^{1,2}, and Deming Liu^{1,2}

¹Next Generation Internet Access Laboratory (NGIA), School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan 430074, China.

²Wuhan National Laboratory for Optoelectronics (WNLO), Wuhan 430073, China
[e-mail: jacquesmbous@gmail.com]

*Corresponding author: Jacques Mbous

*Received June 27, 2018; revised October 16, 2018; accepted November 29, 2018;
published June 30, 2019*

Abstract

Global data center IP traffic is expected to reach 20.6 zettabytes (ZB) by the end of 2021. Intra-data center networks (Intra-DCN) will account for 71.5% of the data center traffic flow and will be the largest portion of the traffic. The understanding of traffic distribution in Intra-DCN is still sketchy. It causes significant amount of bandwidth to go unutilized, and creates avoidable choke points. Conventional transport protocols such as Optical Packet Switching (OPS) and Optical Burst Switching (OBS) allow a one-sided view of the traffic flow in the network. This therefore causes disjointed and uncoordinated decision-making at each node. For effective resource planning, there is the need to consider joining the distributed with centralized management which anticipates the system's needs and regulates the entire network. Methods derived from Kalman filters have proved effective in planning road networks. Considering the network available bandwidth as data transport highways, we propose an intelligent enhanced SDN concept applied to OBS architecture. A management plane (MP) is added to conventional control (CP) and data planes (DP). The MP assembles the traffic spatio-temporal parameters from ingress nodes, uses Kalman filtering prediction-based algorithm to estimate traffic demand. Prior to packets arrival at edges nodes, it regularly forwards updates of resources allocation to CPs. Simulations were done on a hybrid scheme (1+1) and on the centralized OBS. The results demonstrated that the proposition decreases the packet loss ratio. It also improves network latency and throughput—up to 84 and 51%, respectively, versus the traditional scheme.

Keywords: Artificial Intelligence, Hierarchical Networks, Optical Burst Switching, Optical Networks, Resource Allocation, Traffic Estimate.

This work was supported in part by the National Natural Science Foundation of China under Grant 61331010, Grant 61722108 and the 863 High Technology Plan of China under Grant 2015AA016904.

1. Introduction

The high demand and perpetual growth of bandwidth driven by social networks, cloud computing, online gaming, and bandwidth-hungry video services constrain modern communication protocols. Internet extended from thousands of professional users to billions of consumers. This expansion of connected machines resulted in an explosion of traffic flow, that was still north-south (users to data centers traffic). The future of data center IP traffic flow includes intra-data center traffic (71.5%), inter-data center traffic (14.9%), and data center to user traffic (13.6%) [1]. More bandwidth has been allocated to help data center networks cope with this rapid evolution. This improves the capacity but not the efficiency. Effective data analysis combined with intelligent routing decisions and transport strategies are required to solve the explosion of data processing within data centers [2 - 4].

Moreover, due to the abundance of optical bandwidth and low power consumption, data center network architectures are being upgraded to all-optical-networks (AON). This change comes with the development of new transport strategies that fit the main optical domain. Much attention has however been allocated to design optical counterparts of existing electronic strategies such as Optical Circuit Switching (OCS) [5] and Optical Packet Switching (OPS) [6], with mixed results.

Optical Burst Switching (OBS) [7] is a hybrid electronic and optical transport strategy that exists on distributed and centralized architecture. It highlights the advanced logic of electronic components together with the low power consumption, large bandwidth, and high transmission speed of optical devices. However, both OBS architectures statically allocate resources based on the aggregation algorithm. The choice of the aggregation parameter influences the waiting time required for incoming packets in the assembly unit. The three popular techniques (timer, burst-length, mixed) [8], set a fixed time and burst-length parameter. Despite positive aspects, these three techniques tend to be static and reactive. They induce static waiting time and minimum burst-length regardless of the present demand per destination. These procedures lead to an unbalanced use of wavelengths, which creates congested points in the network. Central management can theoretically decrease the waiting time spent by incoming data in the assembly unit, since it enables a global view of the network. However, it requires a longer time to process all data when it is used continuously.

Artificial intelligence techniques can be of great support in the planning of Intra-DCN. Their implementation has allowed an optimization of infrastructures use. However, machine learning techniques still produce a fairly high error rate, raising questions on the use of other methods, including statistical methods. Kalman filtering is a statistical technique which has proven to be effective in tracking positioning and in road networks, where it showed abilities to plan in advance the behavior of vehicles even in some infrequent cases.

In this work, we use Kalman filter-based algorithm to evaluate in advance short-term traffic demand in each direction of an Intra-DCN. The assessment is then involved in centralized and coordinated allocation of all network resources (prior to the arrival of traffic at each node), taking into account the current traffic demand per destination. It is important to note that this does not require any additional waiting time at the edge nodes, the processing being done in background. Our proposition limits bursts waiting time in the assembly unit. Subsequently, we will implement this new planning on two hybrids (centralized and distributed) OBS architectures and one purely centralized OBS. Section 3 gives more details about the proposed architecture and operation.

2. Related Work

Optical burst switching (OBS) [9] is a candidate to carry future Internet traffic. In OBS networks, packets are aggregated at electronic ingress nodes and forwarded as bursts to egress nodes. The OBS uses burst control packets (BCPs) to improve resource utilization, relieve the use of buffering, and reduce costs. Scheduling algorithms such as LAUC-VF [10] and Min-SV [11] estimate the arrival time of bursts from BCPs and reserve network resources from the estimated start time. Sometimes, data bursts do not arrive at intermediate nodes at the estimated time, and these bursts might be stored in optical caches, which increases the use of buffering. Traditional OBS is a distributed-based OBS schemes, it does not have a comprehensive view of the network. In the distributed OBS scheme (see Fig. 1), there is no single point where all of the traffic flow can be viewed at any moment. To address the traditional OBS shortcomings, the time-domain wavelength interleaved networks (TWIN) protocol [12] proposes a centralized scheduling technique where resources can be allocated based on traffic demands. The TWIN uses a central controller to schedule incoming traffic and reduces packet contention. However, even in its most updated versions [13,14], TWIN allocation of sub-wavelengths resources is quasi-static. It is slow to adapt to the change in traffic behavior. TWIN abilities to learn incoming traffic behavior are still sketchy—this can induce latency due to significant waiting time in the assembly unit and unequal load sharing.

Artificial intelligence methods advocate more dynamic concepts to solve this issue [15-17]. Recent studies favor a predictive approach to traffic distribution in order to anticipate its behavior and thus improve the global performance of the network. In [18], Li et al. use the wavelet transform and artificial neural network (ANN) to improve congestion in inter-data center networks. They make predictions on sublink information, elephant flow, and manage to limit the error rate by 5 to 30% compared to existing methods. In [19], Cao et al. perform a statistical analysis of the traffic received by the virtual machines (VMs) and deploy a resource allocation strategy based on the usage history of each VM. Using the Autoregressive Integrated Moving Average (ARIMA) model, the authors are able to increase the bandwidth utilization rate in a cloud data center. In [20], Alvizu et al. propose a routing algorithm for software-defined mobile carrier networks. They use ANN to effectively predict traffic behavior, improve routing decisions and lower power consumption by up to 31% compared to existing standards. In [21], X. Cao et al. propose a mixed prediction model using Convolution Neural Networks (CNN) to study the spatial characteristics of traffic and the Gated Recurrent Unit (GRU), for the temporal factor. The authors manage to improve the error rate, up to 14.3% compared to these methods taken independently.

In Intra-DCN, input data used to evaluate the future behavior of the system (and therefore the allocation of resources) are incomplete. They reflect the measurement of the state of the system during well-defined time intervals. However, according to real-world experiments carried out by T. Benson et al. [31-33], the traffic distribution in Intra-DCN varies greatly over time, highlighting the difficulty to predict the long-term behavior of traffic in Intra-DCN. It partly explains the high error rate of machine learning methods, for both supervised and unsupervised techniques. On the other hand, methods based on Kalman filters have proved effective in areas where the data to be processed presents a great deal of uncertainty. In intelligent transportation systems (ITS), they demonstrated strength in tracking spatio-temporal characteristics of input variables. In [22], Wang et al. use data collected from the ring road of Amsterdam city, in a real-time and large-scale environment. The authors are able to predict 30 min in advance, the occupancy rate of motorways, with a marginal error rate. In addition, the filling rate of on-ramps, leading to these motorways is also predicted with accuracy. In [23], Yuan et al. use Ensemble Kalman Filter to predict traffic in Dutch roads. The authors are able to make accurate estimations of traffic flow compared to other stochastic methods. Their algorithm is able to successfully predict some of unusual scenarios. In [24], Mir et al. emphasized on the

prediction of the speed of motorists. They use the Kalman filter technique to predict the real-time speed of vehicles.

If we consider the available bandwidth, CPU and memory as communication resource pool, then Kalman filtering technique can be used for resource allocation problems in data centers. In [25], Jain et al. use dual Kalman filter to efficiently allocate communication resources, including network bandwidth and memory. The authors are able to highlight the high efficiency of Kalman filters in stream management, especially in terms of algorithmic complexity and their applications in various environments. In [26], Kalyvianaki et al. propose an adaptive resource allocation mechanism in data center environment. The authors use Kalman filtering to efficiently allocate CPU resources to virtual machines (VM). The results are highly effective in dealing with uncertainties in the distribution of data over time. From applications in wireless sensor networks and spacecraft position estimates [27,28], it is established that the technique increases its performance with a large amount of available data, as is the case in Intra-DCN. Our proposition uses Kalman Filtering prediction-based algorithm to make short-term traffic destination estimates. We aim to quickly adapt to change of traffic behavior and improve network congestion in Intra-DCN. The bandwidth use will not be addressed in this paper.

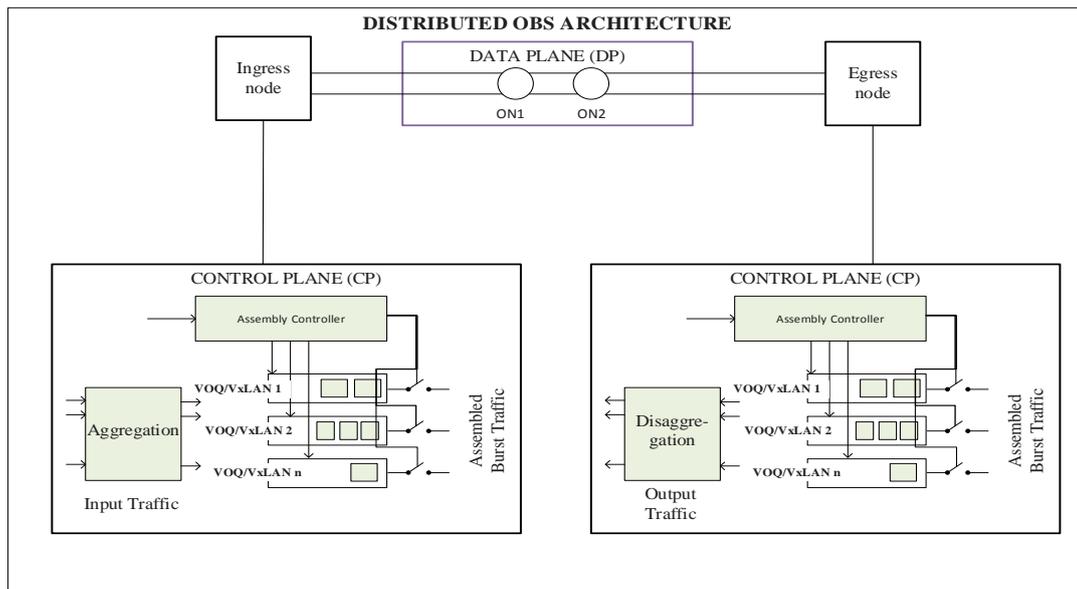


Fig. 1. Distributed OBS Architecture

3. Traffic Prediction

3.1 Proposed Framework

The proposition is a SDN-based network architecture. The proposed model has a global view of network traffic in the entire network. It evaluates traffic distribution to make accurate prediction of future flow destinations. It coordinates resources allocation throughout the network, reducing network congestion (see Fig. 3). Each distributed node only receives relevant updates. Updates are sent at fixed or varied time interval. The remainder of this section describes in more detail the operation of each layer (see Fig. 2).

Ingress traffic are assembled in the assembly unit of the control plane (CP), based on logical destinations—Virtual eXtensible Local Area Network (VXLAN) [29], with the objective to be allocated available resources such that they can efficiently reach their physical destinations. Aggregated bursts will then follow a pre-scheduled framework sent by the management plane (MP). Two frameworks are considered in this proposition, including hybrid and centralized scheme. For the hybrid implementation, packets arriving at ingress nodes—and which are not scheduled to leave for the maximum waiting time—can be redirected to special channels allocated to traditional OBS. They will be dropped when the new centralized framework is running alone.

The MP is built on a predefined network node. This is a layer that records and processes information from the CP, such as the destination and length of each burst. This processing is performed in parallel with the operation of the other layers, which has the advantage of not inducing any additional waiting time at the arrival of the next incoming bursts in a given node. The MP evaluates the directional delay for each traffic flow and uses the Kalman filter-based algorithm to estimate the optimal amount of resources to allocate in each direction for a time cycle. It computes the percentage of full paths to be allocated per intended burst destinations for fixed or varied time periods. A resource allocation frame is built and only sent to CP if it satisfies the update condition. Each distributed node only receives relevant information with respect to its directly connected neighbors. New updates overwrite the current ones at the end of a time cycle.

Update condition: If the period of time T is fixed, then the timer of T should expire before any new updates are sent to CP. When T is variable, new updates are sent when the traditional OBS resource utilization rate (hybrid systems) or packet loss ratio (centralized systems), has reached a predefined threshold (system health).

The Data Plane (DP) will transfer data in the optical domain. It receives aggregated bursts and guarantees fast forwarding without any buffering along the path.

Here, the system health is defined as a measure of the ability of the proposed model to efficiently forward network traffic using the current MP configurations. System health is only evaluated when the time period (T) is variable. Two different measures of system health are being defined. In the hybrid architecture, the reconfigurations are immediately required from the MP if the links allocated to traditional OBS are in high demand (protection line). Second, the system health is measured using the packet loss ratio for the newly centralized architecture.

3.2 Kalman Filtering-based Algorithm

In this section, the stochastic system is used to evaluate the optimal or near-optimal prediction. Comprehensive estimations are formulated through linear and time-invariant systems. Kalman filtering prediction algorithm, which is well-known for its efficiency in data fusion is used to extract short-term patterns. Dynamic input data will be logical traffic destinations and burst lengths that are parameters used to target the resource allocation problem.

We assume to have knowledge of the number of logical end points. The maximum burst-length will also be set during the experiments. Here, we present the optimal estimate in one destination. The same procedure is repeated in the MP to include all of the logical destinations.

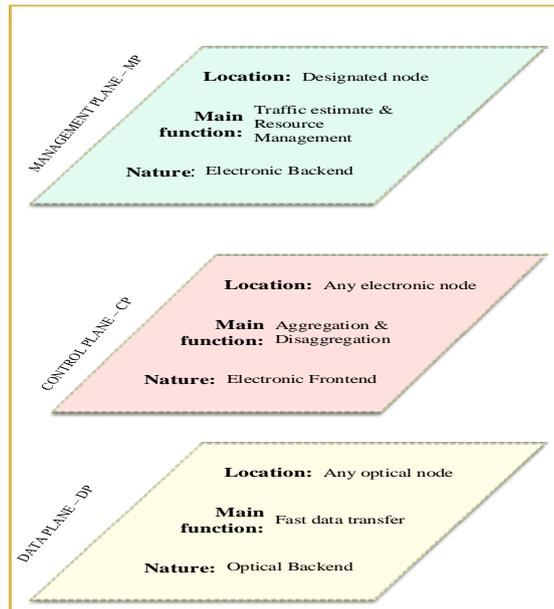


Fig. 2. Layers of Information Processing

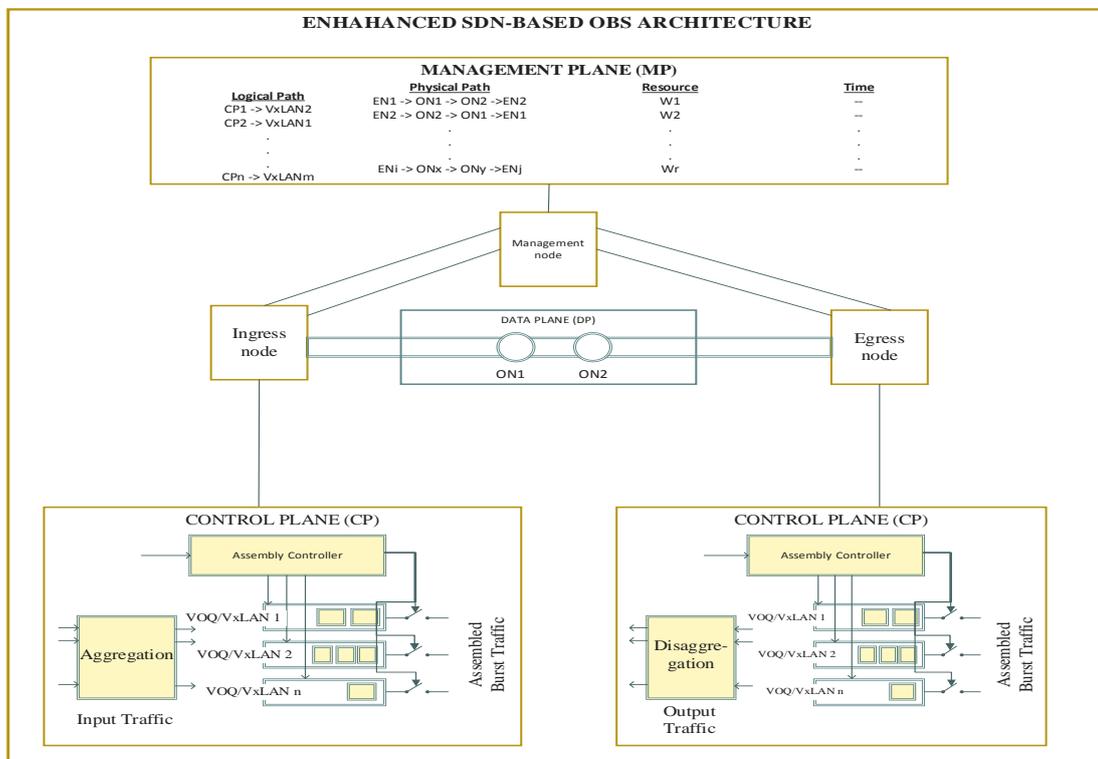


Fig. 3. Enhanced SDN-based OBS Architecture

A linear and discrete time-invariant system is represented by:

$$\begin{cases} x_{t+1} = Ax_t + Bu_t + \Omega\alpha_t \\ y_t = Cx_t + Du_t + \beta_t \end{cases} \quad (1)$$

where A, B, C, D and Ω are known constant matrices, respectively, with $1 \leq m, p, q \leq n$. $\{u(t)\}$, a known deterministic control input sequence of m -vectors. $\{\alpha(t)\}$ and $\{\beta(t)\}$ are system and observation noise sequences, respectively, with zero-mean Gaussian white noise, and respective variance, $Var(\alpha(t)) = Q_\alpha$, $Var(\beta(t)) = R_\beta$, positive definite matrices.

We will assume that: $E(g_0\alpha_t^T) = 0$ and $E(g_0\beta_t^T) = 0$.

Equation (1) describes the linear deterministic and stochastic system which can be divided into purely deterministic and stochastic systems.

The stochastic system is represented by:

$$\begin{cases} g_{t+1} = Ag_t + \Omega\alpha_t \\ m_t = Cg_t + \beta_t \end{cases} \quad (2)$$

While the deterministic system is described as:

$$\begin{cases} f_{t+1} = Af_t + Bu_t \\ s_t = Cf_t + Du_t \end{cases} \quad (3)$$

$$\text{with } x_t = f_t + g_t \quad (4)$$

$$\text{and } y_t = s_t + m_t \quad (5)$$

and the transition equation:

$$f_t = (A_{t-1} \dots A_0)x_0 + \sum_{i=1}^t (A_i \dots A_{i-1}) B_{i-1} u_{i-1} \quad (6)$$

We use (2) to derive the stochastic optimal estimate \hat{g}_t of g_t so that $\hat{g}_{t-1} = \hat{g}_{t-1|t-1}$.

From data vector $\begin{bmatrix} z_0 \\ \vdots \\ z_k \end{bmatrix}$,

the linear stochastic system can be written as:

$$\bar{z}_k = \Pi_{t,k} g_t + \bar{\epsilon}_{-t,k}$$

$$\text{with } \Pi_{t,k} = \begin{bmatrix} C_0 \Psi_{0,t} \\ \vdots \\ C_k \Psi_{k,t} \end{bmatrix} \quad \text{and } \bar{\epsilon}_{-t,k} = \begin{bmatrix} \bar{\epsilon}_{-t,0} \\ \vdots \\ \bar{\epsilon}_{-t,k} \end{bmatrix}.$$

$\Psi_{l,t}$ are transition matrices defined as:

$$\Psi_{l,t} = \begin{cases} A_{l-1} \dots A_t & \text{if } l > t, \\ I & \text{if } l = t \end{cases}$$

For $l < t$, $\Psi_{l,t} = \Psi_{t,l}^{-1}$

then, $\varepsilon_{t,l} = \beta_{-l} - C_l \sum_{i=l+1}^t \Psi_{li} \Omega_{i-1} \alpha_{i-1}$ (7)

thus, $g_t = \Psi_{t,l} g_l + \sum_{i=l+1}^t \Psi_{ti} \Omega_{i-1} \alpha_{i-1}$

it follows: $g_t = \Psi_{t,l} g_l + \sum_{i=l+1}^t \Psi_{ji} \Omega_{i-1} \alpha_{i-1}$

leading to:

$$\begin{aligned} \Pi_{t,k} g_t + \bar{\varepsilon}_{-t,k} &= \begin{bmatrix} C_0 \Psi_{0t} \\ \vdots \\ C_k \Psi_{kt} \end{bmatrix} g_t + \begin{bmatrix} \beta_0 - C_0 \sum_{i=1}^t \Psi_{0i} \Omega_{i-1} \alpha_{i-1} \\ \vdots \\ \beta_{-k} - C_k \sum_{i=k+1}^t \Psi_{ki} \Omega_{i-1} \alpha_{i-1} \end{bmatrix} \\ &= \begin{bmatrix} C_0 g_0 + \beta_0 \\ \vdots \\ C_k g_k + \beta_k \end{bmatrix} \\ &= \begin{bmatrix} z_0 \\ \vdots \\ z_k \end{bmatrix} = \bar{z}_k \end{aligned}$$

Through this experiment, we consider the recursive formula only including data information about: $\hat{g}_{t-1} = \hat{g}_{t-1|t-1}$ to make predictions of $\hat{g}_t = \hat{g}_{t|t}$.

We can derive the recursive formula, which is the real-time estimate of the stochastic problem

$$\begin{cases} \hat{g}_{t|t} = \hat{g}_{t|t-1} + G_t (z_t - C_t \hat{g}_{t|t-1}) \\ \hat{g}_{t|t-1} = A_{j-1} \hat{g}_{t-1|t-1} \end{cases} \quad (8)$$

Where G_t are Kalman gain matrices, and \hat{g}_0 will be an unbiased estimated of the initial state g_0 . We can now compute the optimal least-squares estimate $\hat{g}_{t|k}$ of g_t .

By defining the weight $W_{t,k} = (\text{var}(\bar{\varepsilon}_{-t,k}))^{-1}$,

thus,

$$\begin{aligned} W^{-1}_{t,t-1} &= \begin{bmatrix} R_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R_{t-1} \end{bmatrix} + \text{Var} \begin{bmatrix} C_0 \sum_{i=1}^t \Psi_{0i} \Omega_{i-1} \alpha_{i-1} \\ \vdots \\ C_k \sum_{i=k+1}^t \Psi_{ki} \Omega_{i-1} \alpha_{i-1} \end{bmatrix} \\ W^{-1}_{t,t-1} &= \begin{bmatrix} W^{-1}_{k,k-1} & 0 \\ 0 & R_k \end{bmatrix} \end{aligned}$$

then,

$$\hat{g}_{t|k} = (\Pi_{t,k}^T W_{t,k} \Pi_{t,k})^{-1} \Pi_{t,k}^T W_{t,k} \bar{z}_t$$

Let's now define:

$$\begin{aligned} G_t &= (\Pi_{t,t-1}^T W_{t,t-1} \Pi_{t,t-1} + C_t^T R_t^{-1} C_t)^{-1} C_t^T R_t^{-1} \\ &= (\Pi_{t,t-1}^T W_{t,t} \Pi_{t,t})^{-1} C_t^T R_t^{-1} \end{aligned}$$

We can then write:

$$\begin{aligned} \hat{g}_{t|t} &= \hat{g}_{t|t-1} + G_t (z_t - C_t \hat{g}_{t|t-1}), \text{ with} \\ \hat{g}_{t|t-1} &= A_{t-1} \hat{g}_{t-1|t-1} \end{aligned}$$

We can now compute:

$$\begin{aligned} G_t &= P_{t,t-1} C_t^T (C_t P_{t,t-1} C_t^T + R_t)^{-1} \quad (9) \\ P_{t,t} &= (\Pi_{t,t}^T W_{t,t} \Pi_{t,t})^{-1} \\ P_{t,t-1} &= (\Pi_{t,t-1}^T W_{t,t-1} \Pi_{t,t-1})^{-1} \end{aligned}$$

It follows that:

$$\begin{aligned} G_t &= P_{t,t-1} C_t^T (C_t P_{t,t-1} C_t^T + R_t)^{-1} \\ P_{t,t} &= (I - G_t C_t) P_{t,t-1}, \\ P_{t,t-1} &= A_{t-1} P_{t-1,t-1} A_{t-1}^T + \Omega_{t-1} Q_{t-1} \Omega_{t-1}^T \end{aligned}$$

Moreover,

$$\begin{aligned} P_{t,t-1} &= E(g_t - \hat{g}_{t|t-1})(g_t - \hat{g}_{t|t-1})^T \\ &= \text{Var}(g_t - \hat{g}_{t|t-1}) \\ P_{t,t} &= E(g_t - \hat{g}_{t|t})(g_t - \hat{g}_{t|t})^T \\ &= \text{Var}(g_t - \hat{g}_{t|t}) \quad (10) \\ P_{0,0} &= \text{Var}(g_0) \end{aligned}$$

We can summarize this procedure as follows:

$$\begin{aligned} P_{0,0} &= \text{Var}(g_0) \\ P_{t,t-1} &= A_{t-1} P_{t-1,t-1} A_{t-1}^T + \Omega_{t-1} Q_{t-1} \Omega_{t-1}^T \\ G_t &= P_{t,t-1} C_t^T (C_t P_{t,t-1} C_t^T + R_t)^{-1} \\ P_{t,t} &= (I - G_t C_t) P_{t,t-1} \\ \hat{g}_{0|0} &= E(x_0) \\ \hat{x}_{t|t-1} &= A_{t-1} \hat{x}_{t-1|t-1} + B_{t-1} u_{t-1} \\ \hat{x}_{t|t} &= \hat{x}_{t|t-1} + G_t (z_t - D_t u_t - C_t \hat{x}_{t|t-1}) \\ &t = 1, 2, 3, \dots \end{aligned}$$

Algorithm 1: Resource Allocation Process - Management Plane

1: Initialize the clock

2: Set a time cycle for T

3: If {T is fixed},

While {T > 0}

No updates will be sent; step 3 will be repeated.

else {/T is variable/}, call the function `systemhealth ()`

% The function `systemhealth ()` permanently evaluates the current state of wavelength use of traditional OBS (hybrid scheme) and compares it to a preset threshold or it instead evaluates the packet loss ratio (centralized scheme). The function will always return a Boolean%.

If {`systemhealth ()` returns 0},

No Updates will be sent; step 3 will be repeated.

else {/systemhealth () returns 1/}

4: Call the function `kalmanpredict ()`

It evaluates the ratio of traffic per VXLAN destination.

5: Resource Allocation

Here, the algorithm assigns resources at predefined central point and distributes information to every electronic ingress nodes in respect to their needs.

Algorithm 1 describes the process of dynamic resource allocation in the MP.

First, a mode of operation is selected for the time cycle T. The mode will be defined as fixed or variable.

Second, updates are made differently depending on whether the selected time cycle is fixed or variable. In the case of a fixed mode, an initial value is assigned to T. The value of the time cycle will be decremented, and the CP will not receive any updates from the MP until this value is equal to 0 (the time cycle has elapsed). If, however, the chosen mode is variable, an initial value is no longer necessary. The reception of the updates is conditioned by the result of the system health function.

Third, the function `system health` works as follows: if the protection line utilization rate or packet loss ratio, reaches a certain threshold set in advance, then it is time for the MP to provide CP with further updates. Hybrid systems (1 + 1) use the protection line utilization rate to assess the overall health of the system. In purely centralized systems (not having a protection line), the packet loss ratio is used. This function returns a boolean. If the boolean is 1, new updates can be sent to the specific CP. If the boolean is 0, the current updates are considered to be optimal.

Finally, the updates sent by the MP to the CP of a specific node, derive from the Kalman filter estimates made from the input data, such as the destination and the length of the bursts. The MP build a resource allocation framework for each node in the network.

4. Simulation Results and Discussions

4.1 Simulation Platform

Hybrid and Centralized OBS were evaluated here. The results were compared to the traditional distributed scheme. In Hybrid OBS (1+1), the Centralized OBS is implemented as main line for traffic forwarding, while the distributed OBS will be used for protection. To maintain unscheduled full paths in the hybrid architecture, running the traditional OBS as protection can acknowledge the inability of the proposed algorithm to reach 100% successful traffic estimation. The expectation over hybrid architectures is to provide the maximum QoS with a flexible network architecture.

The C++ programming language was used to build our network model in OMNET++ simulation platform. Three-tier data center network architecture with 100 Gbps links, is used (see Fig. 4).

We assumed that most servers can run different applications at the same moment. VMs running the same application are grouped together following the concept of VXLAN, which can theoretically support up to 16 million logical networks. A network link originates at the electronic ingress nodes and ends at the electronic egress nodes (except intra-rack paths). This is the name of the full path. There is one path per wavelength from the ingress node A to the egress node B. The topic of sub-wavelengths is out the scope of this paper.

The Software-defined concept here refers to the use of the 3 levels of traffic management, namely the DP for transmission by optical fibers, the CP for aggregation / disaggregation of data in the assembly units and finally the MP for the prediction and coordinated planning of resources. In this work, this concept does not refer to the use of tools such as Openflow.

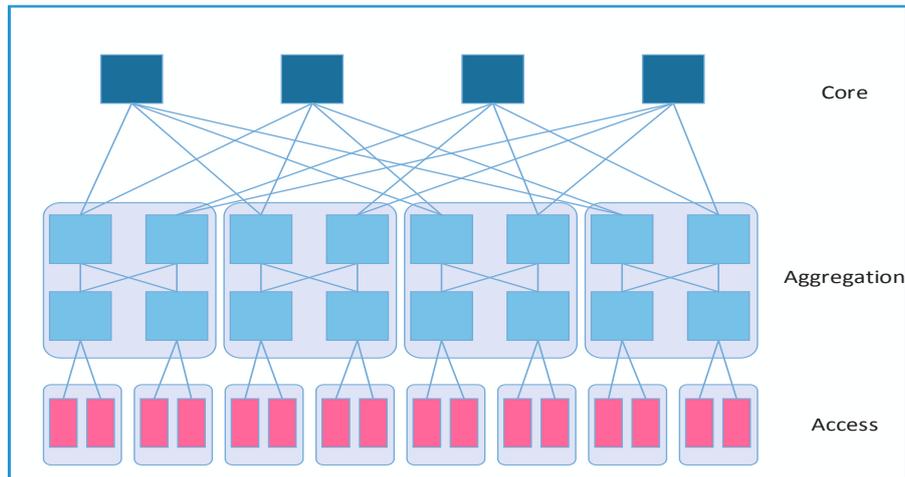


Fig. 4. Data Center Network Architecture

A- Traffic distribution

We acknowledge the superiority of real-world dataset over simulation data. but their great disparity from one data center to another hardly guarantees their absolute reliability for this experiment. However, according to Benson et al. [33], Weibull and lognormal traffic distribution might better describe data center traffic. Lognormal distribution will be used here to simulate traffic flow, since it is simple to be put together with Kalman filters. If X is

log normally distributed, then $Y=\ln(X)$ is normally (Gaussian) distributed.

B- Simulation Parameters

Our simulation model includes a set of 100 racks—each includes 50 servers. Buffers of edge nodes can store up to 2500 packets of different sizes. The aggregated value of incoming packets is set to 1 μ s. Packets are aggregated to have a maximum burst-length of 40 KB. Bursts that exceed the predefined 40 KB length have not been considered because they can easily be divided, forwarded, and reassembled at the receiving node following the TCP process. The maximum waiting time of bursts in the assembly unit is 2 μ s, which is quite low compared to tens or hundreds of microseconds for the traditional OBS. Bursts that exceeds the maximum waiting time are immediately redirected to the protection channel (traditional OBS). Maximum electronic switching time per burst is 1 μ s while the optical domain only produces an overhead of 300 ns.

From ingress to egress nodes, each link capacity is set to 100 Gbps for ease of management and implementation. TOR switches are connected to intra-rack servers. We allocate 1 Gbps to each VM.

The speed of computation, in microseconds, is far less than the period T of sending updates from MP to CP (in seconds). Our algorithm runs in background and does not directly induce additional waiting time. However, algorithm complexity is $O \log (n)$, which keeps the performance of the algorithm for higher scale of data.

Table 1. KEY SIMULATION PARAMETERS

Symbol	Quantity	Value
R	Racks	100
S	Servers/Rack	50
TOR	TOR Switches	100
VM	VM Bandwidth	1 Gbps
FP	Full Path Bandwidth	100 Gbps
MPO	MP Overhead	100 ms
T_{elec}	Electronic Switching Time	1 μ s
T_{opt}	Optical Switching Time	200 ns
T_{bcp}	BCP Processing Time-Traditional OBS	1 μ s
T_{prop}	Propagation Time	2 μ s
BS	Burst Size	2500 Packets
L	Length	40 KB
W	Wavelength/Link	80
T_{queue}	Maximum Waiting Time (Queue) in CP	2 μ s
T_{ass}	Assembly Time	40 μ s

4.1 Results and Discussions

To evaluate the performance of the proposed framework, we compared the throughput, latency, and packet loss ratio aspects with the traditional OBS network. Our results include three scenarios of OBS implementation. The first is a hybrid architecture including the traditional model and the new model with the proportion 80/20—here, 80% of the available resources are dedicated to the proposed model and 20% were allocated to the traditional OBS. Second, we modify the proportions to 90/10 with 90% for the proposed model and 10% for the traditional OBS. Third, experiments are carried out on the pure proposed model (100).

A- Latency

We evaluate the average full path (from electronic ingress to electronic egress nodes) network latency of the proposition. Full path traffic flow includes inter-rack (TOR switch-to-TOR switch), rack-to-router, and router-to-router traffic flow—this creates more latency and congestion. Although intra-rack traffic represents a significant amount of traffic flow, this therefore causes a delay of only nanoseconds.

We used the following equation to evaluate the network delay:

$$DELAY = T_{prop} + T_{wait} + T_{elec} + T_{opt}$$

where T_{prop} is the full path propagation delay of the burst, and T_{wait} is the accumulated waiting time induced by CP at every node that includes burst assembly/disassembly time, maximum burst departure waiting time, and packet processing delay. T_{elec} and T_{opt} are electronic and optical switching times, respectively. It should be noticed that MP does not add additional delay to the network, since it runs in the backend. New updates overwrite the old ones.

The simulation results for network latency are presented in **Fig. 5. Fig. 5.a** which shows an 80/20 channel allocation. The proposed hybrid model performed better than the traditional OBS in every aspect excluding two scenarios. First, when $T=100$ s, the graph was unstable. Sometimes it performed very well, and sometimes it performed very badly compared to the traditional OBS. This zigzag behavior is due to the time period set for network updates, which was too long. This confirmed the hazardous estimates when T is relatively long. When the traffic distribution suddenly changes, we wait too long before resource reallocation. The situation worsened when T increased. Second, simulations have limitations in the hybrid schemes with a fixed T . The initial resource allocation in the MP uses an equal cost multipath (ECMP), but the lognormal distribution generates unequally distributed traffic. Time is needed to close the gap. The hybrid architecture with a varied T benefits from the traditional OBS strength at the start. The traffic that waits more than $2 \mu\text{s}$ is immediately reallocated to the traditional OBS. This triggers reallocation of resources if resource utilization in the protection line reaches a predefined threshold of 50% of available resources—this is a very low rate and provides ample room for further improvements.

In **Fig. 5.c**, the proposed centralized model is implemented alone; these results were surprisingly very close to the scenario of 90/10 (**Fig. 5.b**), which is globally the best in terms of network latency. We think that this strange behavior might come from the accuracy of traffic estimates in the MP, which limits the use of a protection channel. This centralized framework does not have a protection line. It only relies on the accuracy of traffic estimates, which seems to be very precise in this case. Studies about the confidence interval of traffic estimates are beyond the scope of this paper.

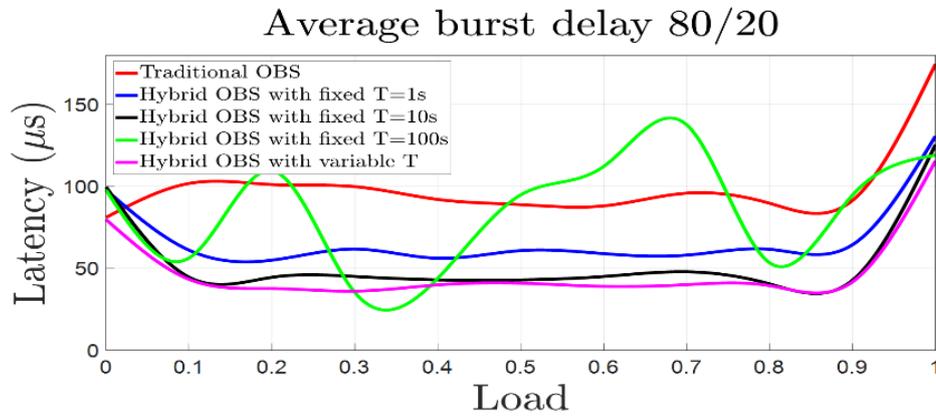


Fig. 5.a. Average Burst Latency - hybrid OBS 80/20

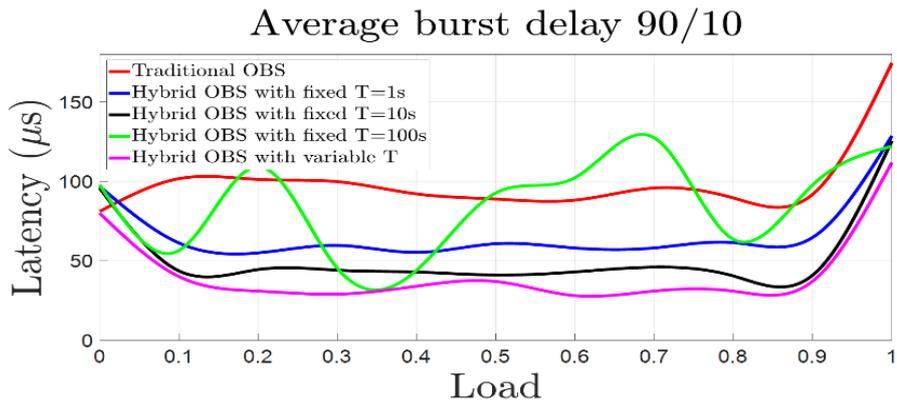


Fig. 5.b. Average burst delay – hybrid OBS 90/10

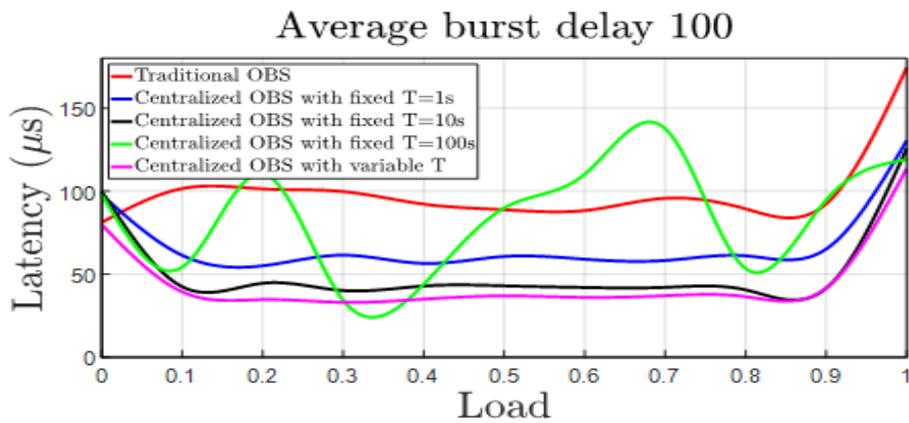


Fig. 5.c. Average burst delay - centralized OBS

B- System Health and Packet Loss Ratio

System health measures the ability of the proposed OBS model to efficiently carry traffic using the current MP configurations. It is only used when T , the time period for updates, is variable. For 80/20 and 90/10 designs, the system health is given by the traditional OBS full path occupation level, and by the packet loss ratio for the 100% experiment.

Fig. 6.a and **Fig. 6.b** show that the average utilization rate often stays very low—even less than 10% for a load < 0.9. For variable T , although network updates are triggered by link utilization at instant peaks, we noticed fewer network reconfigurations needed for the MP compared to the fixed cycle where reconfigurations are periodically planned at a ratio of 1 to 3.

Fig. 7.a, **Fig. 7.b**, and **Fig. 7.c** present simulation results of the packet loss ratio for each scenario (80/20, 90/10, 100): **Fig. 7.a** and **Fig. 7.b** present the network performance. **Fig. 7.c** is a measure of the average system health of a 100 design. In **Fig. 7.a** and **Fig. 7.b**, every scenario of the proposed model performed better than the traditional OBS. For $T=100$ s, the performance is generally better than the traditional model while also being unstable. Combining these results with the instability of the aforementioned scenario, we established that $T \geq 100$ seconds is too long to generate accurate resource allocation.

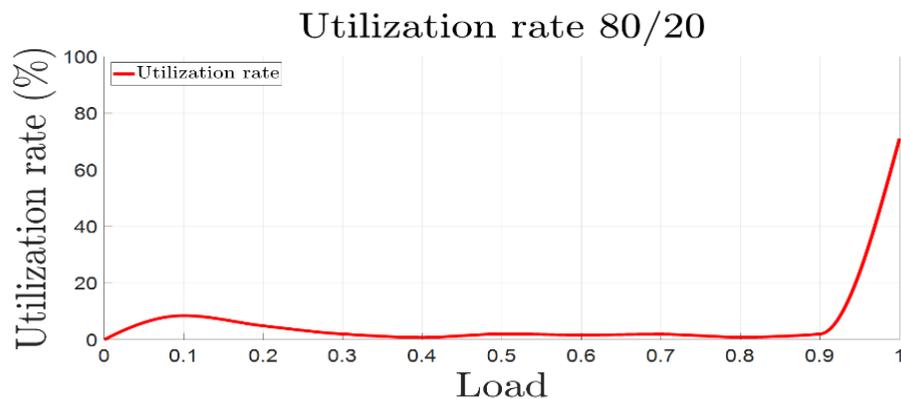


Fig. 6.a. Utilization Rate of Traditional OBS – 80/20

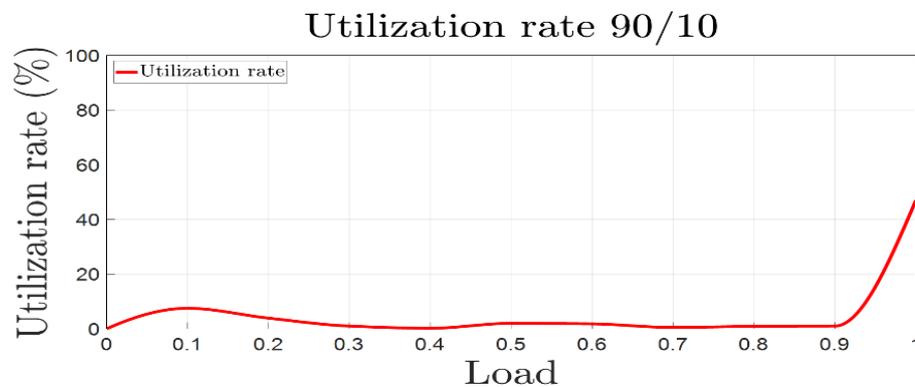


Fig. 6.b. Utilization Rate of Traditional OBS – 90/10

Fig. 7.c shows the good performance of the proposition framework compared to the traditional OBS. However, early packet loss was recorded. We think that earlier packet loss occurs because of sudden traffic change—especially at higher loads. This packet loss challenges the confidence of the estimates under all traffic conditions. We set the threshold of the packet loss ratio at 10^{-2} , which suggests that instant peaks can trigger network updates.

For loads ≤ 0.8 , the proposed framework presents no packet loss because of the absence of BCP contention and unsuccessful estimates of burst arrival. Higher loads—especially at instant peaks of the lognormal distribution—could support resource reallocation on demand within an acceptable period of time. It is interesting to point out that packet loss in 80/20 and 90/10 designs is mainly associated with the running of a traditional OBS (control packet contention and unmatched burst arrival). The opposite packet loss was found in the 100 model resulting from packets dropped in the waiting queue in CP.

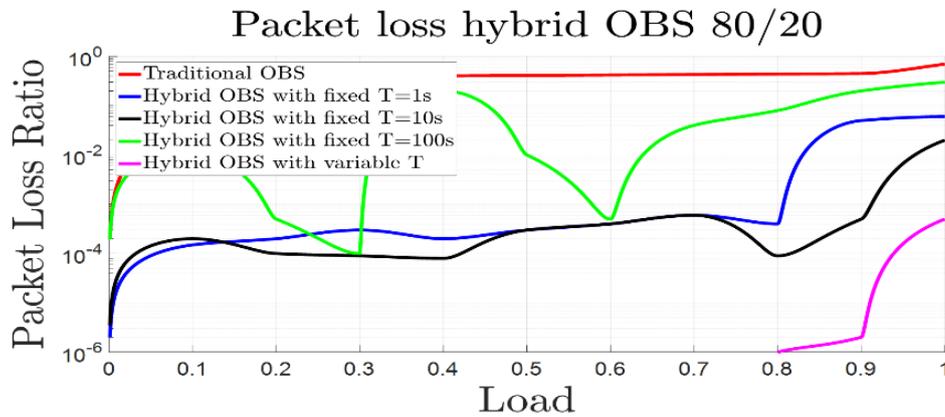


Fig. 7.a. Packet loss ratio – hybrid OBS 80/20

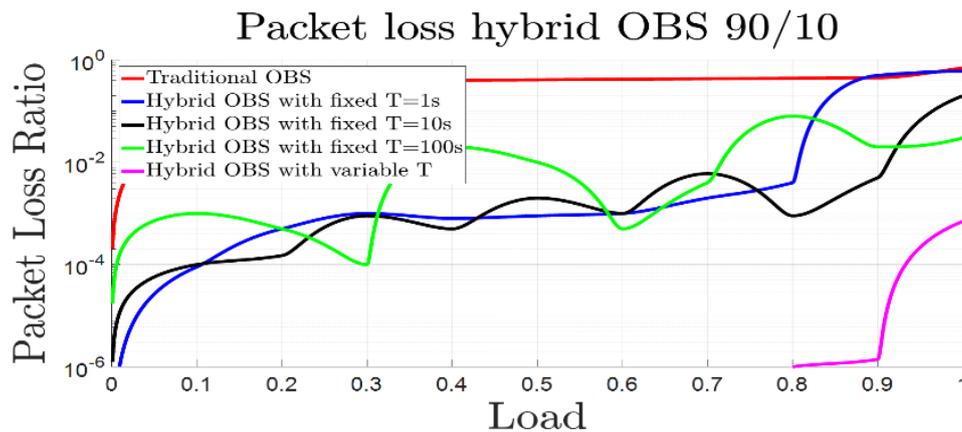


Fig. 7.b. Packet loss ratio – hybrid OBS 90/10

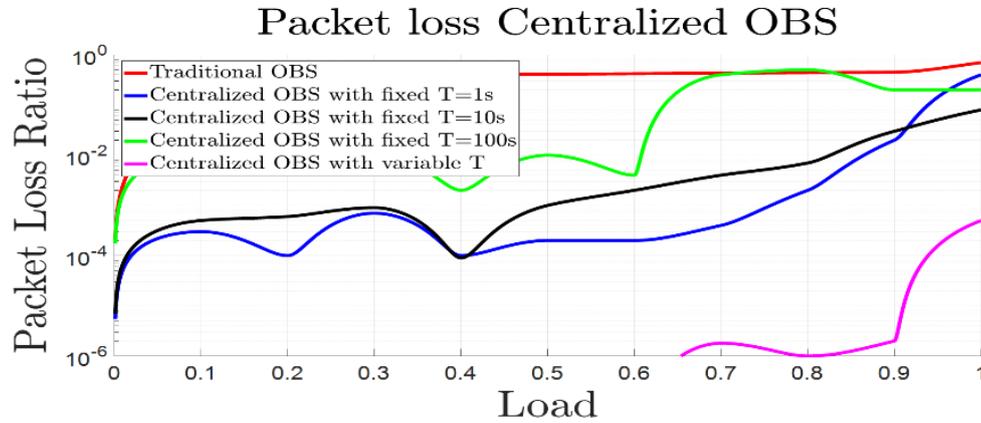


Fig. 7.c. Packet loss ratio – Centralized OBS

C-Throughput

We evaluate the throughput of each full path in the network as follows:

$$Th_{full\ path} = \frac{Total_{bits}}{T_{simulation}}$$

Here, $Total_{bits}$ is the total number of bits delivered successfully in each full path during the simulation time $T_{simulation}$. Links are assumed to be 100 Gbps bandwidth for a fully subscribed network, and throughput in each physical segment is the same. Fig. 8.a, Fig. 8.b and Fig. 8.c present the average throughput of the global network.

The proposition performs better than the traditional model except for the case of fixed $T = 100s$, which produced similar results to the traditional OBS. In the scenarios of variable T or fixed $T = 10s$, the average throughput was improved by 51 and 45%, respectively (Fig. 8.b and Fig. 8.c).

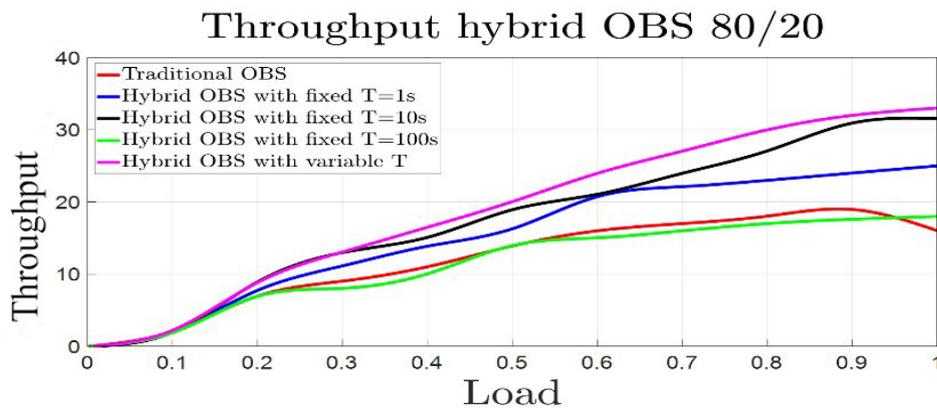


Fig. 8.a. Throughput – hybrid OBS 80/20

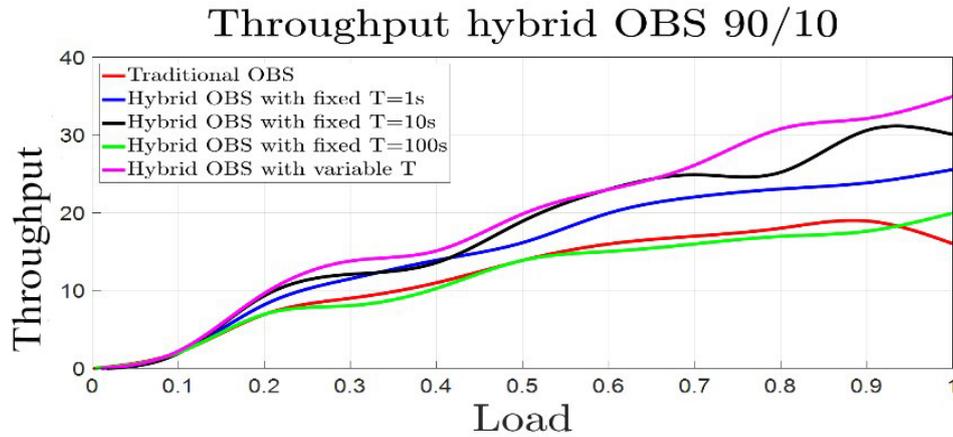


Fig. 8.b. Throughput – hybrid OBS 90/10

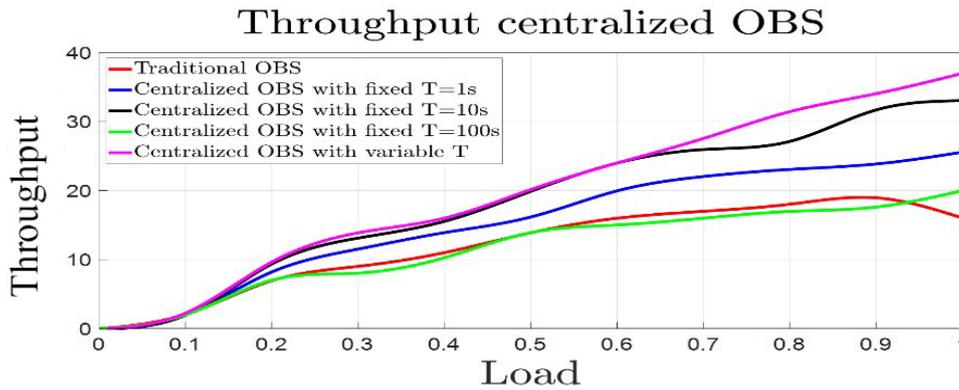


Fig. 8.c. Throughput – Centralized OBS

Although the hybrid systems (1 + 1), with variable T, have the best performances studied, they also induce an additional degree of complexity by the introduction of the System Health function. In the case of the same hybrid systems with fixed T, this factor disappears. The important parameter in this case becomes the choice of T. This choice is strongly dependent on the great variability of the distribution of data over time. If the distribution evolves too abruptly for quite short periods of time, a relatively small choice of T would be appropriate for the experiment. On the other hand, the more the distribution will be stable, a relatively long time cycle T will produce better results. It should be borne in mind that the primary scope of this work is to be able to use it for distributions that change abruptly over time, due to the fact that the nature of incoming traffic in Intra-DCN is not yet accurately described.

5. Conclusion

Here, we used an artificial intelligence technique to improve the global network latency of an OBS system in Intra-DCN. We proposed a SDN-based three layers' architecture. The

proposition includes DP, CP, and MP. A Kalman filtering prediction-based algorithm was used in the MP to estimate short-term traffic horizons. These predictions are used for resource allocation and node scheduling prior to data arrival at electronic ingress nodes. Our algorithm runs in backend, which did not induce any additional waiting time to the bursts in CP. Previous updates remain valid at each ingress node, until new updates can satisfy the necessary and sufficient condition. Our simulations have included three scenarios that was namely two hybrids and one centralized architecture. Hybrid architecture with 90/10 channel distribution tends to be the most suitable for Intra-DCN traffic due to good traffic estimates and fine balance between main and protection line. The time cycle T , for OBS nodes to receive updates from the MP was a main point of discussion. The hybrid systems with fixed time period $T=10$ s have the best trade-off in terms of global network performance and algorithm complexity, while variable T has the best performances in term of packet loss ratio, average network latency, and throughput. Finally, the simulations have confirmed the unpredictable behavior of Intra-DCN for relatively long-term estimates. For $T \geq 100$ seconds, estimates tend to be imprecise. This frequently provided a graph with a significant difference in the amplitude between two close points.

Further researches would include extended Kalman filtering, and bring Kalman filtering together with a machine learning technique. We would then investigate average bandwidth use and the waiting time in the assembly unit. More datasets would also be considered.

References

- [1] Cisco global cloud index: forecast and methodology, 2016-2021 white Paper. [Article \(CrossRef Link\)](#).
- [2] G. Bell, T. Hey, and A. Szalay, "Beyond the data deluge," *Science*, vol. 323, no. 5919, pp. 1297-1298, March, 2009. [Article \(CrossRef Link\)](#).
- [3] Q. Yu, C. Han, L. Bai, J. Choi, X. Chen, "Low-complexity multiuser detection in millimeter-wave systems based on opportunistic hybrid beamforming," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 10129-10133, October, 2018. [Article \(CrossRef Link\)](#).
- [4] C. Zhang, W. Zhang, "Spectrum sharing for drone networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 136-144, November, 2017. [Article \(CrossRef Link\)](#).
- [5] K. J. Barker, A. Benner, R. Hoare, H. Hoisie, A. K. Jones, D. K. Kerbyson, D. Li, R. Melhem, R. Rajamony, E. Schenfeld, S. Shao, C. Stunkel, and P. Walker, "On the feasibility of optical circuit switching for high performance computing systems," in *Proc. of 2005 ACM/IEEE Conf. on Supercomputing*, pp. 16, November 12-18, 2005. [Article \(CrossRef Link\)](#).
- [6] M. J. O'Mahony, D. Simeonidou, and D. K. Hunter, "The application of optical packet switching in future communication networks," *IEEE Communication Magazine*, vol. 39, no. 3, pp. 128-135, March, 2001. [Article \(CrossRef Link\)](#).
- [7] C. Qiao, M. Yoo, "Optical burst switching: a new paradigm for an optical Internet," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69-84, 1999. [Article \(CrossRef Link\)](#).
- [8] J. P. Jue, W. -H. Yang, Y. -C. Kim, and Q. Zhang, "Optical packet and burst switched networks: a review," *IET Communications*, vol. 3, no. 3, pp. 334-352, March, 2009. [Article \(CrossRef Link\)](#).
- [9] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: a new area in optical networking research," *IEEE Network*, vol. 18, no. 3, pp. 16-23, May-June, 2004. [Article \(CrossRef Link\)](#).
- [10] Y. Xiong, M. Vandenhouste, and H. C. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1838-1851, October, 2000. [Article \(CrossRef Link\)](#).

- [11] J. Xu, C. Qiao, J. Li, and G. Xu, "Efficient burst scheduling algorithms in optical burst-switched networks using geometric techniques," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1796–1811, November, 2004. [Article \(CrossRef Link\)](#).
- [12] K. Ross, N. Bambos, K. Kumaran, I. Saniee, and I. Widjaja, "Scheduling bursts in time-domain wavelength interleaved networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1441–1451, November, 2003. [Article \(CrossRef Link\)](#).
- [13] G. Cazzaniga, C. Hermsmeyer, I. Saniee, and I. Widjaja, "A new perspective on burst-switched optical networks," *Bell Labs Technical Journal*, vol. 18, no. 3, pp. 111-131, December, 2013. [Article \(CrossRef Link\)](#).
- [14] A. Triki, I. Popescu, A. Gravey, X. Cao, T. Tsuritani, and P. Gravey, "TWIN as future-proof optical transport technology for next generation metro networks," in *Proc. of 2016 IEEE 17th Int. Conf. on HPSR*, pp. 87-92, June 14-17, 2016. [Article \(CrossRef Link\)](#).
- [15] D. Zibar, M. Piels, R. Jones, and C. G. Schaeffer, "Machine learning techniques in optical communication," *IEEE/OSA Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1442–1452, March, 2016. [Article \(CrossRef Link\)](#).
- [16] M. Glick, H. Rastegarfar, "Scheduling and control in hybrid data centers," in *Proc. of 2017 Photonics Society Summer Topical Meeting Series (SUM)*, pp. 115-116, July 10-12, 2017. [Article \(CrossRef Link\)](#).
- [17] J. Mata, I. de Miguel, R. J. Duran, N. Merayo, S. K. Singh, A. Jukan, and M. Chamania, "Artificial intelligence (AI) methods in optical networks: a comprehensive survey," *Elsevier Optical Switching and Networking Journal*, vol. 28, pp. 43-57, April, 2018. [Article \(CrossRef Link\)](#).
- [18] Y. Li, H. Liu, W. Yang, D. Wu, X. Wang, and W. Xu, "Predicting inter-data-center network traffic using elephant flow and sublink information," *IEEE Transactions on Network and Service Management*, vol. 13, no. 4, pp. 782–792, December, 2016. [Article \(CrossRef Link\)](#).
- [19] J. Cao, X. Zhu, F. Dong, B. Liu, Z. Ma, H. Min, "Time series based bandwidth allocation cloud strategy in cloud datacenter," in *Proc. of 2016 Advanced Cloud and Big Data*, pp. 228-233, August 13-16, 2016. [Article \(CrossRef Link\)](#).
- [20] R. Alvizu, S. Troia, G. Maier, A. Pettavina, "Matheuristic with machine-learning-based-prediction for software-defined mobile metro-core networks," *IEEE/OSA Journal of Optical Communication and Networking*, vol. 9, no. 9, pp. D19-D30, September, 2017. [Article \(CrossRef Link\)](#).
- [21] X. Cao, Y. Zhong, Y. Zhou, J. Wang, C. Zhu, and W. Zhang, "Interactive temporal recurrent convolution network for traffic prediction in data centers," *IEEE Access*, vol. 6, pp. 5276–5289, December, 2017. [Article \(CrossRef Link\)](#).
- [22] Y. Wang, J. H. van Schuppen, and J. Vrancken, "Prediction of traffic flow at the boundary of a motorway network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 214–227, February, 2014. [Article \(CrossRef Link\)](#).
- [23] Y. Yuan, F. Scholten, H. van Lint, "Energy traffic state estimation and prediction based on the ensemble Kalman filter with a fast implementation and localized deterministic scheme," in *Proc. of 2015 IEEE 18th International Conference on ITSC*, pp. 477-482, September 15-18, 2015. [Article \(CrossRef Link\)](#).
- [24] Z. H. Mir, F. Filali, "An adaptive Kalman filter based traffic prediction algorithm for urban road network," in *Proc. of 2016 12th Int. Conf. on IIT*, pp. 1-6, November 28-30, 2016. [Article \(CrossRef Link\)](#).
- [25] A. Jain, E. Y. Chang, Y. Wang, "Adaptive stream resource management using Kalman filters," in *Proc. of the 2004 ACM SIGMOD Int. Conf. on Management of Data*, pp. 11-22, June 13-18, 2004. [Article \(CrossRef Link\)](#).
- [26] E. Kalyvianaki, T. Charalambous, S. Hand, "Self-adaptive and self-configured CPU resource provisioning for virtualized servers using Kalman filters," in *Proc. of the 6th Int. Conf. on Autonomic Computing*, pp. 117-126, June 15-19, 2009. [Article \(CrossRef Link\)](#).

- [27] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453-1464, September, 2004. [Article \(CrossRef Link\)](#).
- [28] E. J. Lefferts, F.L. Markley, and M.D. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417-429, September-October, 1982.
- [29] M. Mahalingam, D. Butt, K. Duda, P. Argawal, L. Kreeger, T. Sridhar, M. Bursell, and C. Wright, "Virtual eXtensible local area network (VXLAN): a framework for overlaying virtualized layer 2 networks over layer 3 networks," *RFC 7348*, August, 2014. [Article \(CrossRef Link\)](#).
- [30] C. Nguyen Le Tan, C. Klein, and E. Elmroth, "Location-aware load prediction in edge data centers," in *Proc. of 2017 2nd Int. Conf. on FMEC*, pp. 25-31, May 8-11, 2017. [Article \(CrossRef Link\)](#).
- [31] [Article \(CrossRef Link\)](#).
- [32] T. Benson, A. Anand, A. Akella, and M. Zhang, "Understanding data center traffic characteristics," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, pp. 92-99, January, 2010. [Article \(CrossRef Link\)](#).
- [33] T. Benson, A. Anand, A. Akella, and M. Zhang, "The case for fine-grained traffic engineering in data centers," in *Proc. of INM/WREN*, pp. 2-2, April 28-30, 2010. [Article \(CrossRef Link\)](#).
- [34] T. Benson, A. Akella, and D. Maltz, "Network traffic characteristics of data centers in the wild," in *Proc. of the 10th ACM SIGCOMM Conf. on Internet Measurement*, pp. 267-280, November 01-30, 2010. [Article \(CrossRef Link\)](#).
- [35] H. Rastegarfar, M. Glick, N. Viljoen, M. Yang, J. Wissinger, L. LaComb, and N. Peyghambarian, "TCP flow classification and bandwidth aggregation in optically interconnected data center networks," *IEEE/OSA Journal of Optical Communication and Networking*, vol. 8, no. 10, pp. 777-786, October, 2016. [Article \(CrossRef Link\)](#).
- [36] P. N. Ji, "Hybrid optical-electrical data center networks," in *Proc. of Advanced Photonic Congress, Photonic Network and Devices*, July 18-20, 2016. [Article \(CrossRef Link\)](#).
- [37] C. Kachris, K. Kanonakis, and I. Tomkos, "Optical interconnection networks in data centers: recent trends and future challenges," *IEEE Communication Magazine*, vol. 51, no. 9, pp. 39-45, September, 2013. [Article \(CrossRef Link\)](#).
- [38] M. Yoo and C. Qiao, "Just-enough-time (JET): a high speed protocol for bursty traffic in optical networks," in *Proc. of 1997 Digest of the IEEE/LEOS Summer Topical Meeting: Vertical-Cavity Lasers/Technologies for a Global Information Infrastructure/WDM Components Technology/Advanced Semiconductor Lasers and Application*, pp. 26-27, August 11-13, 1997. [Article \(CrossRef Link\)](#).
- [39] J. Y. Wei, R. I. McFarland Jr, "Just-in-time signaling for WDM optical burst switching networks," in *IEEE/OSA Journal of Lightwave Technology*, vol. 18, no. 12, pp. 2019-2037, December, 2000. [Article \(CrossRef Link\)](#).
- [40] I. Widjaja, "Performance analysis of burst admission-control protocols," *IEE Proceedings - Communications*, vol. 142, no. 1, pp. 7-14, February, 1995. [Article \(CrossRef Link\)](#).
- [41] Y. Xiong, M. Vandenhoute, and H. Cankaya, "Design and analysis of optical burst-switched networks," in *Proc. of SPIE, All-Optical Networking: Architecture, Control, and Management Issues*, vol. 3843, pp. 112-119, August 25, 1999. [Article \(CrossRef Link\)](#).
- [42] J. S. Turner, "Terabit burst switching," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 3-16, March, 1999. [Article \(CrossRef Link\)](#).
- [43] G. Weichenberg, V. Chan, and M. Médard, "Design and analysis of optical flow-switched networks," *IEEE/OSA Journal of Optical Communication and Networking*, vol. 1, no. 3, pp. B.81-B.97, August, 2009. [Article \(CrossRef Link\)](#).
- [44] C. K. Chui, and G. Chen, "Kalman Filtering with Real-Time Applications," 5th Edition, Springer, Stanford, 2017. [Article \(CrossRef Link\)](#).
- [45] T. Truong-Huu, M. Gurusamy, and S. T. Girisankar, "Dynamic flow scheduling with uncertain flow duration in optical data centers," *IEEE Access*, vol. 5, pp. 11200-11214, June, 2017. [Article \(CrossRef Link\)](#).



Jacques Mbous received his M.Sc in the department of Electronic and Information Engineering from Huazhong University of Science and Technology, (HUST), Wuhan, China, in 2013. He is currently PhD candidate in the Next Generation Internet Access laboratory (NGIA), School of Optical and Electronic Information of HUST. His current research interests include flexible optical network, computational intelligence and data center networks.



Tao Jiang received the B.E. degree from the State Key Laboratory of Integrated Service Network, School of Telecommunication Engineering, Xidian University, China, 2015. He is currently a Ph.D student in Huazhong University of Science and Technology, China, since 2015. His current interests include flexible optical network, access network and data center network.



Ming Tang (SM'11) received his B. Eng. from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2001, and the Ph.D. degree from Nanyang Technological University (NTU), Singapore, in 2005. His postdoctoral research in Network Technology Research Centre (NTRC) was focused on the optical fiber amplifiers, high-power fiber lasers, nonlinear fiber optics, and all-optical signal processing. From February 2009, he was with Tera-photonics group led by Prof. Hiromasa Ito in RIKEN, Sendai, Japan, as a Research Scientist conducting research on terahertz-wave generation, detection, and application using nonlinear optical technologies. Since March 2011, he has been a Professor in the School of Optical and Electronic Information, Wuhan National Laboratory for Optoelectronics, HUST, Wuhan, China. His current research interests are concerned with optical fiber based linear and nonlinear effects for communication and sensing applications. Dr. Tang has been a member of IEEE Photonics Society since 2001, he became the senior member of IEEE since 2011 and he is also a member of OSA.



Songnian Fu received the B.Sc. and M.Sc. degrees from Xiamen University, Xiamen, China, in 1998 and 2001, respectively, and the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2005. From 2005 to 2011, he was with the Network Technology Research Center, Nanyang Technological University, Singapore, as a Research Fellow. Since February 2011, he has been a Professor in the School of Optical and Electronic Information and the Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, China. His research interests include fiber-optic transmission and microwave photonics.



Deming Liu received his Ph.D. degree in the department of Electronic and Information Engineering from Huazhong University of Science and Technology, (HUST), Wuhan, China, in 1999. From 1999 to 2000, he was with the Network Technology Research Center, Nanyang Technological University, Singapore, as a Research Fellow. He has been a Professor in the School of Optical and Electronic Information and the Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, China. His research interests include optical access and wireless access.