A Real-Time Admission Control and Planning of OLS Networks

Yassine Khlifi and Noureddine Boudigua
Network Telecommunications and security (CNAS) laboratory, University of November 7th
Carthage, Tunisia
Corresponding author E-mail Address: nab@supcom.rnu.tn

***Abstract*— Optical Label Switching (OLS) has been proposed as new paradigm to provide optical resource provision and support of Quality of service (QoS) satisfaction for application with various constraints. Admission Control and planning are critical tasks affecting OLS networks, since the network design and data operational costs are fundamental criterion in networking platform and operation. In this paper, we introduce a novel technique for network resource and traffic management based on real-time handling of the network capability and traffic requirements. This technique handles signaling tasks, admission control and contention resolution. We develop a scalable admission control and planning formulation approach based on buffering and transmission capabilities, and traffic requirements. Finally, simulation experiments are performed in order to validate our proposal.**

I. INTRODUCTION

The rapid grow of Internet traffic, and the emergence of novel applications and services has triggered new requirements in terms of resources allocation and quality of service (QoS) support. Optical Label Switching (OLS) has been proposed as a suitable switching paradigm for providing fast switching capability, resource provisioning and QoS support [1]. Admission control and planning are critical tasks in OLS network, since network design and efficient data transmission costs are fundamental design criterion which influenced the choice of the networking infrastructure and operation. Admission control can be implemented for preserving traffic requirements and achieving a high level of resource utilization. Whereas planning can be build in order to dimension and optimize the network resource as well as to efficiently manage the traffic parameters. Presently, the network resource dimensioning and optimization formulation, in presence of an advanced call admission protocol (CAC) scheme, is an important design issue in OLS networks. Such formulation can be applied for optimizing network resource allocation and utilization in terms of buffering and transmission capacity, and traffic parameters.

To our knowledge, little work has considered providing admission control, and networks resource dimensioning and optimization in optical switched QoS-oriented traffic parameters. These methods provided a static approach using mean value analysis (MVA) [3] in which the traffic requirements and network resource availability are managed through the use of mean value of traffic and system parameters. Other works (e.g. [5]) have addressed virtual topology and traffic grooming optimization, where traditional optimization methods are used. However, all these methods do not take into account dynamic real-time variation of system status and traffic requirements as well as cannot achieve a suitable decision for resource allocation and utilization, and high level of QoS satisfaction.

In this work, we address the issue of admission control and planning of OLS networks based on a dynamic real-time handling of traffic parameters and network resource capability. We propose a scalable approach and formalize it using the incoming traffic requirements in terms of the delay and loss. The proposed approach considers a dynamic real-time management of node architecture and traffic characteristics, such as the time variation of output traffic rate and buffering capacity, at each core node of the considered path, as well as the input traffic parameters and wavelengths capability, along the transmission lightpath. We also develop an analytic model for formulating separately the admission control problem, and resource dimensioning and optimization problem by means of a conservation law [4] and queuing network model. Finally, simulation study has been conducted in which a comparative study of real-time approach and MVA approach is performed in order to validate the efficiency of our novel approach.

The remaining part of this paper is organized as follows. Section 2 briefly studies the basic concepts of the OLS network architecture and schemes as well as the admission control and planning problem. Section 3 develops a theoretical model that can help to design the appropriate admission control and planning problem formulation using a conservation law [4] and queuing network model. Section 4 discusses the validation of the proposed approach through a simulation work. Section 5 concludes the paper.

II. OLS NETWORKS ARCHITECTURE AND SCHEMES

We describe here an OLS network architecture and, the dynamic signaling and contention resolution schemes as well as the admission control and network planning problem.

A. OLS core node and network architectures

An OLS network consists of a collection of edge and core nodes. The network uses the core node architecture proposed and analyzed in [2]. The nodes are interconnected with DWDM links. Core nodes transfer IP traffic from an input port to its destination based on new label switching framework built during a signaling step [3] containing traffic requirements and real traffic parameters. The core node performs signaling protocol, and adopts dynamic contention resolution using wavelength converters and ODL buffers. The node architecture is typically composed of M input and M output ports. Each port consists of several channels which can handle with w wavelengths and a set of multiplexers and demultiplexers. For a complete description of this architecture, the reader can refer to [2]. The main components of this architecture are: the input processing unit (IPU), switch control unit (SCU), switch fabric unit (SFU), waiting unit (WU), and output processing unit (OPU).

B. Dynamic OLS signaling scheme

In OLS network, the transmission process is preceded by the lightpath, label switching path (LSP), setup that defines the route between the ingress node and egress node. Dynamic Signaling protocol as it is proposed in [3] uses two new label frameworks: label request and label mapping. The first label is created by the ingress edge and sent to the egress node. Its structure contains the requested traffic requirements in terms of delay and loss. The second label is built and sent back to the ingress node in order to confirm the LSP establishment and update the switching tables of the concerned LSP-core nodes. In this case, the traffic is transmitted over the network core using the established LSP. IP packets have a variable-length, and the switching label which insures that the switching process is completed ahead of the incoming traffic packets. This label carries the switching wavelength value, the traffic requirements and real QoS values processed along the related LSP-core nodes.

C. Dynamic contention resolution scheme

The dynamic contention resolution is made based on the wavelength conversion and ODL buffering (in that order). The QoS-related aspect shows that the choice of the packet to convert or to delay, in the contention case, is taken based on traffic parameter states on the entire network. For a complete description of the dynamic contention resolution, the reader can refer to [3]. When contention occurs, the core node compares the difference between contract parameters and real parameters in terms of delay and loss of contended packets. The authorized packet is directed to its original output port whereas the contended packet is converted to an alternative wavelength of the established LSP to avoid conflict. If no wavelength is available, the packet is stored in one among the available ODL buffers. According to the feed-backward buffer utilization located in WU, a delayed packet traverses an ODL, and re-enters the switch. If no ODL is available the packet is dropped. If contention is still occurring, a packet may be buffered several times without exceeding the requested blocking delay of the related traffic.

D. Admission control and planning problem

OLS networks aim at improving resource allocation and providing an efficient resource utilization to support various applications constraints. Dynamic signaling and contention resolution schemes provide an over-dimensioning of resource allocation and inefficient use of the allocated resource through the use of MVA approach [3]. To alleviate these problems, we introduce a novel dimensioning and optimization approach using dynamic real-time variation of resource capability and a suitable call admission control protocol for providing an optimized signaling cost, an efficient resource allocation, and a better traffic management. To implement our approach, we need to consider the traffic parameters, the estimation of the arrival probability of the traffic i delayed j times during the interval I_k, where $I_k \in [k.T, (k+1).T]$. k is defined as a multiple number of packets length (i.e. $k \in [0, (m_i-1)]$, m_i represents the end-to-end delay constraint used to resolve contention and T denotes the slot duration or packet length ($0 \leq j \leq m_i$). We also consider the estimation of the probability of the resource availability in terms of transmission and buffering availability. The decision-making for admission control and network resource planning is made based on the estimated parameters and traffic requirements along the related LSP.

III. ANALYTIC MODEL FOR ADMISSION CONTROL AND PLANNING

We present hereinafter the analytical model developed for admission control and planning of the built lightpath.

A. Modeling

To model an LSP, we consider a core node which handles N traffic types labeled 0,1,…N-1 that are composed of variable-length traffic packets. Let T, d, and w denote, mean packet length, buffering capacity of the node and transmission capability available at each output port, respectively. Each traffic i is assumed to require a maximum packet loss, L_{max} and a maximum packet blocking delay, D_{max}. Let $PT^k_{i,j}$, for $0 \leq j \leq m_i$, represents the traffic sub-type consisting of packets of type i which have been buffered j times during the interval I_k as it is defined in the previous subsection. $PT^k_{i,0}$ represents initial arriving packets of traffic i at the node. In addition, we assume that arriving packets of type i addressed to a specific output port of the node follow Poisson process with rate $\lambda^k_{i,0}$. Packets are transmitted with needed QoS, real QoS status and integrate the adequate traffic sub-types. Finally, $B^k_{i,j}$ and $F^k_{i,j}$ denote the blocking probability due to the lack of wavelengths at output port and the blocking probability due to the lack of available ODL in buffering unit for $PT^k_{i,j}$ traffic type.

The proposed model is an open queuing network system, as it is defined and analyzed in [2, 3], is composed of two queues. Queue 1 (say q1), which represents the output port transmission queue, has an $M/D/k/k$ preemptive dynamic type. Queue 2 (say q2), which represents the waiting unit, has an M/D/d/d type. The system is assumed to handle

$N*(m_i+1)$ traffic sub-type. Let us consider now the path followed by a packet through the queuing network system. The newly arriving packet (i.e. $PT^k_{i,0}$) can be switched immediately or be moved to the WU. In the first case, a q1 is allocated to this packet during a service time T (i.e. packet length), and then the packet leaves the system. In the second case, the packet is sent to q2 where it can also be served or be dropped. When the packet is served, the q2 is allocated during a deterministic service time T. Once its service at q2 is completed, it moves to q1 as a packet of the traffic of sub-type $PT^k_{i,1}$. Let note that each time the packet returns to q1, its traffic sub-type is updated by incrementing it according to its stay duration j ($\in PT^k_{i,j+1}$). Because buffering threshold is fixed to m_i, a packet of sub-type $PT^k_{i,mi}$ which cannot seize one server at q1 will be dropped. In addition, each packet transmitted from q1 to q2, where $j<m_i$, needs to be buffered. If no ODL is available, the packet is dropped.

Now, we consider a lightpath consisting of n core nodes. Based on the assumptions made for a node, it becomes easy to model a lightpath. The arriving packet of traffic i, $PT_{i,j}$ at an output port follows the similar behaviors as described for a core node. Let now $PT^{1,k}_{i,j}$, $0 \leq i \leq N-1$, $0 \leq j \leq m_i$, $0 \leq l \leq n-1$ represents the traffic sub-type at the 1st core node (i.e. node 0) consisting of packets of type *i* which have been delayed *j* times during I_k. $PT^{0,k}_{i,0}$ represents a newly arriving packets of type *i* at the 1st node during I_k. Moreover, we assume that arriving packets of type *i* that are addressed to a specific output port of each core node follows also a Poisson process with rate $\lambda^{0,k}_{i,0}$. We denote the blocking probability due to the lack of wavelengths at output port by $B^{n,k}_{i,j}$ and the blocking probability due to the lack of the available ODL by $F^{n,k}_{i,j}$, $0 \leq i \leq N-1$, $0 \leq j \leq m_i$ for the traffic of sub-type $PT^{n,k}_{i,j}$.

B. *Model analysis*

Let us consider a path traversed by a packet through the queuing network of the established path and suppose that the arriving packet is an element of sub-type $PT^l_{i,0}$, $0 \leq l \leq n-1$. According to the stay duration j ($\in PT^{l,k}_{i,j+1}$) described for one core node, we note that each packet successfully served at a q1 of the each core node, its traffic sub-type is updated according to node number *l* ($\in PT^{l+1,k-j}_{i,j}$). The analysis has been conducted using a conservation law that was initially proposed for the evaluation of burst blocking probabilities in an optical burst switching (OBS) network [4].

To analyze the blocking probabilities of the different traffic sub-types of our analytical model, we follow the analysis based on the considered conservation law. Let $\lambda^k_{i,j}$ denotes the arrival rate of the traffic sub-types $PT^k_{i,j}$, 0≤i≤N-1, 0≤j≤mi during I_k interval. Then, $\lambda^k_{i,j}$ is given by:

$$\lambda^k_{i,j} = \lambda^{k-j}_{i,0} \prod_{l=0}^{j-1} B^{(k-j)+l}_{i,l} \left(1 - F^{(k-j)+l}_{i,l}\right) \qquad (1)$$

$$\text{where } j \in [1, k] \text{ if } k < m_i \text{ Else } j \in [1, k]$$

Let $X^{n,k}_{i,j}$, $X^{n,k}_j$ and *a* denote a random variable which represent the number of packet (of traffic of type i) arrived at a core node n with j delay during I_k and *a* the considered packet number, respectively. In this case, the probability of the considered random variable is given by:

$$P(X^{n,k}_j \leq a) = \left(\frac{\sum_{i=0}^{N-1} \lambda^{n,k}_{i,j}}{\lambda^{0,k-j}_0} \right) \sum_{l=0}^{a} \left(\lambda^{0,k-j}_0\right)^l . \frac{e^{-\lambda^{0,k-j}_{i,0}}}{l!} \qquad (2)$$

$$P(X^{n,k}_{i,j} \leq a) = \left(\frac{\lambda^{n,k}_{i,j}}{\lambda^{0,k-j}_{i,0}} \right) \sum_{l=0}^{a} \left(\lambda^{0,k-j}_{i,0}\right)^l . \frac{e^{-\lambda^{0,k-j}_{i,0}}}{l!} \qquad (3)$$

Let $PL^{n,k}$ *and* $PL_i^{n,k}$ denotes the overall packet loss and the packet loss of traffic i measured at LSP-core node n during I_k. Using the previous probability formulations, $PL^{n,k}$ and $PL_i^{n,k}$ are given by the following expressions:

$$PL^{n,k}_i = \frac{X^{n,k}_{i,m_i} B^{n,k}_{i,m_i} + \sum_{j=0}^{m_i-1} X^{n,k}_{i,j} B^{n,k}_{i,j} F^{n,k}_{i,j}}{\sum_{j=0}^{m_i} X^{0,k-j}_{i,0}} \qquad (4)$$

$$PL^{n,k} = \frac{\sum_{j=0}^{m_i} X^{n,k}_j B^{n,k}_j F^{n,k}_j}{\sum_{j=0}^{m_i} X^{0,k-j}_j} \text{, where } F^{n,k}_{m_i} = 1 \qquad (5)$$

C. *Admission control and planning decision*

Based on collected information and resource capability, each core node of the considered LSP can contribute to accept the incoming traffic and characterize the network resource availability based on the formulations (6) and (7).

$$P\left(PL_i^{n,k} \leq L_{Max}\right) \geq 1 - \varepsilon \qquad (6)$$

$$P\left(PL^{n,k} \leq L_{Max}\right) = f(d) = g(w)$$

$$\begin{cases} Minimize(d) \\ f(d) \geq 1-\varepsilon \\ with\ w\ link\ capacity \end{cases} \qquad \begin{cases} Minimize(w) \\ g(w) \geq 1-\varepsilon \\ with\ d\ buffering\ capacity \end{cases} \qquad (7)$$

$$where\ \varepsilon: system\ parameter$$

Where the probability of $PL^{n,k}$ can be formulated based on two functions *f* and *g* which are given by the following expressions: $f\,(d, k, \lambda^{n,k}_{i,0}, m_i)$ and $g\,(k, d, \lambda^{n,k}_{i,0}, m_i)$.

IV. SIMULATION AND NUMERICAL RESULTS

We present here the implemented simulation model and discuss some of the most important numerical results.

A. *Simulation Model*

In our simulation model, we have found interesting to consider four input traffic types for the arriving packets denoted by T0, T1, T2 and T3. T0 has the lowest traffic loss and transmission delay tolerance. T3 has the highest loss and blocking delay. We assume that the input traffic ratios of packet traffics are 10%, 20%, 30%, and 40% for traffic types

T0, T1, T2, and T3, respectively. All packets have a variable length generated uniformly distributed in the interval [250, 1500] bytes. The following performance metrics have been chosen to resolve the formulated problem: the packet loss mean rate and the average packet-blocking delay. Two input system parameters have been considered: the inter-arrival mean time and traffic load (packet number).

B. *Numerical results*

Figures 1-2 depict the average packet-blocking delay and packets-loss mean rate versus the impact of the inter-arrival mean time. These figures illustrate that, for all traffic types, the increase of the input parameter decreases the output parameters. This can be explained by the fact that the increase of the input parameter increases the time separating two successive packets decreases the traffic charge, which induces the decrease of the risk of contention, and so the output parameters. Figures 3-4 plot the average packet-blocking delay and packets-loss mean rate versus the impact of traffic load. We notice that, for all traffic types, the increase of the input parameter increases the output parameters. This is because the growth of packet number increases the traffic load, which increases the risk of contention and the output parameters. We observe that, the real-time approach performs better than MVA approach and offers a high level of QoS satisfaction in terms of loss and delay, for all traffic types. This is because the proposed approach provides real-time status of system parameters and traffic characteristics which can help each LSP core node to reach a suitable decision on admission control request processing, and resource dimensioning and optimization.

V. CONCLUSION

In this paper, we mainly addressed the issue of admission control and planning for OLS networks. We introduced a new technique for dimensioning and optimization in presence of novel admission control scheme based on dynamic real-time management of traffic requirements and network resource in terms of transmission and buffering capabilities. Based on traffic contract and real traffic parameters, dynamic differentiation is performed trough the use of wavelength conversion, and ODL buffering techniques to resolve the contention. We have also developed a theoretical model for formulating the proposed approach by mean of a conservation law and queuing network model. The simulation results show that the dynamic real-time approach can effectively better reduce the loss and delay compared to the MVA approach.

REFERENCES

[1] S. J. B. Yoo et al, “Optical-label switching based packet routing system with contention resolution capabilities in wavelength, time,.” in Proc. OFC’2002, Anaheim, Mar. 2002.

[2] Y. Khlifi, S. Guemara El Fatmi, and N. Boudriga “Optical Label Switching: Design Issues, Modeling, and Performance Analysis”, ICECS’05, Tunisia, Dec. 2005.

[3] Y. Khlifi, N. Boudriga, and M. S. Obaidat “A QoS-based Scheme for Planning and Dimensioning of Optical Label Switched Networks”, ICC’07, Glasgow, Scotland, June 2007.

[4] Guoping Zeng, Kejie Lu, and Imrich Chlamtac “On the Conservation Law in Optical Burst Switching Networks”, SPECTS, San Diego, pp.124-129, Jul. 2004.

[5] Datta P., Sridharan M., Somani A.K., "A Simulated Annealing Approach for Topology Planning and Evolution of Mesh-Restorable Optical Networks", ONDM-2003, Budapest, Hungary, Feb. 2003.

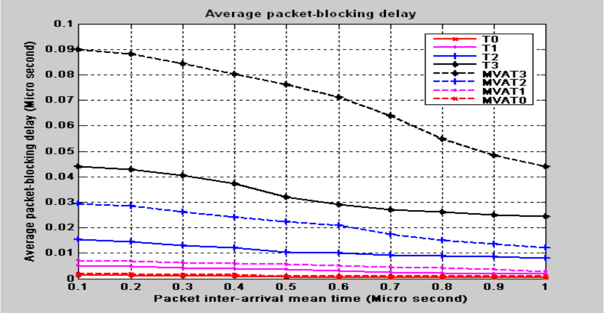

Figure 1: Blocking delay versus inter-arrival mean time

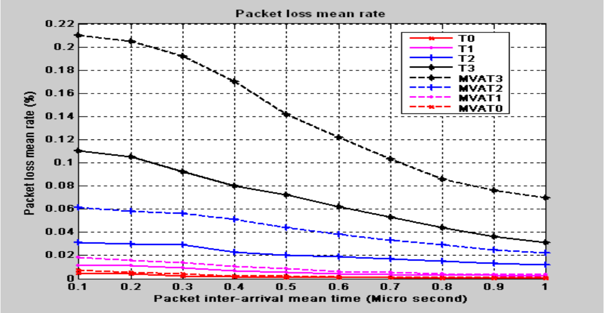

Figure 2: Packets-loss mean rate versus inter-arrival mean time

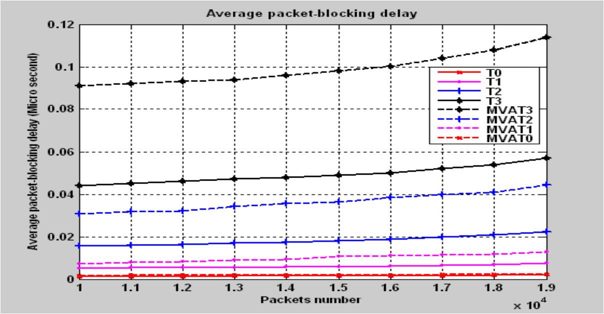

Figure 3: Blocking delay versus traffic load

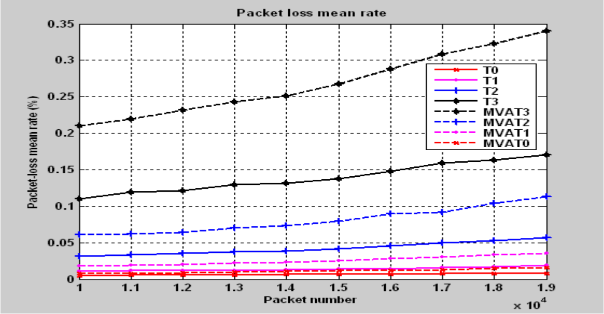

Figure 4: Packets-loss mean rate versus traffic load