Impact of Traffic Delay Tolerance on Elastic Optical Networks Performance

Jaume Comellas, Farzaneh Nourmohammadi, Gabriel Junyent
Optical Communications Group (GCO), Universitat Politècnica de Catalunya (UPC)
UPC Campus Nord D5, Jordi Girona 1, 08034 Barcelona (Spain), Tel: (+34) 93 401 6441
Email: comellas@tsc.upc.edu

ABSTRACT

Elastic Optical Networks (EON) are considered as a valuable solution to enhance optical networks efficiency due to their superior network spectrum use. By breaking the traditional fixed grid of WDM networks, the spectrum assigned to connection demands is adapted to their requirements. On the other hand, connection parameters such as baud-rate or modulation format are dynamically adjusted taking into account the current network conditions. Therefore, EONs fit perfectly with the heterogeneous characteristics of current and future network traffic loads. Nevertheless, when EONs are in operation, they present some weaknesses mainly derived from the spectrum fragmentation generated by the dynamic establishment and torn-down of these diverse connections. The randomness of connections generation results in the blocking of some connections even when the total load supported by the network is not so high. In this work, we analyse how traffic tolerance to delay impacts on the EON performance, and present simulation obtained results when connection demands have no strict delay constraints compared with demands which need to be real-time served.

Keywords: elastic optical network, performance analysis, traffic characterization

1. INTRODUCTION

The Elastic Optical Network (EON) architecture [1] was proposed in response to the huge capacity demand and the diverse granularity needs of current and future Internet traffic. By adjusting the spectrum allocated to the connections' needs (assigning variable sized spectrum bands to them), the EONs clearly improve in terms of spectral efficiency. Depending on its bandwidth requirements, the appropriate amount of optical spectrum is assigned to each connection. Similar to WDM networks, an elastic optical connection must occupy the same portion of spectra between its end nodes, that is, to ensure the so-called spectrum continuity restriction. In addition, the full bandwidth of the connections will be allocated contiguously. This is known as the spectrum contiguity constraint. On the other hand and as it was introduced in [2], the time continuity constraint mandates that the resources are allocated for the complete connections duration.

The routing and spectrum allocation (RSA) problem in EONs has been widely studied, with special emphasis on dynamic (in-operation) network scenarios (see [3] and its references). In a dynamic traffic scenario, connection requests are sent to the EON as they arrive. These requests (source and destination nodes, bandwidth needed and duration) start and finish in a random way and are served or not depending on the network resources status at the time of each request. The network state evolves randomly depending on the connections established and completed. Each time a new request arrives, the RSA algorithm is executed in real-time to decide whether it is possible to allocate the intended connection. Connections which cannot be allocated are blocked, so the bandwidth blocking probability (BBP, ratio of the total requested bandwidth which cannot be allocated) is a key performance parameter in EONs. After analysing EON performance, it is shown that in many situations the usage of the resources available in the network is in general very poor when a low BBP has to be guaranteed [4]. Concretely, it was found that in EONs, the average occupation of spectrum under dynamic conditions is around 50% when BBP=0.01 has to be guaranteed [4]. This is mainly due to spectrum fragmentation whose mitigation requires quite complex processes involving connections reallocation and huge management overhead [5]. On the other hand, randomness of connections source and destination nodes makes that some links are overloaded during some periods (e.g. many connections with same origin and destination generated during a short period) even the average occupation during the whole observation period is below 50%. Therefore, the idea of delaying some connections and allocating them in a future time slot (traffic whose delay constraints are not strict is present in any network), when hopefully the traffic through their paths will be lower, seems to be plausible. In this work, we develop this idea and try to measure the impact of delay tolerant traffic on the performance on EONs. The main objective consists in finding generic trends which give insight into the effects of traffic delay tolerance on network performance. These effects will be evaluated by means of extensive simulations under different conditions.

The remaining of this paper is organized as follows. In Section 2, the main concepts about traffic delay tolerance as well as a review of previous works on the topic are introduced. Section 3 is devoted to the description of the simulator implemented and the results obtained. Main work conclusions are summarized in Section 4.

2. TIME AWARENESS IN EONs ROUTING AND SPECTRUM ASSIGNMENT

As commented before, RSA problem in EONs has been widely studied, with special emphasis on dynamic network scenarios [3, 6]. There, connection arrival and departure processes are random, and the network has to accommodate incoming traffic in real time. The spectral resources allocated to connections are released upon tear down, becoming available for future requests. The randomness in the connections' birth and death processes leads to fragmentation of the spectral resources in the network. Indeed, this randomness leads also to huge differences in the occupation of network resources at different time periods. This effect is shown in Fig. 1, which shows the instantaneous number of existing connections in a typical EON under dynamic traffic. It can be observed how this number shows great variance and reaches values in the range of 720-730 at some instants while at other periods decreases to values below 620. The average number of connections during the whole simulation period is 660. It is likely that connections are rejected (blocked) when the links composing their path are more crowded than usual and it is seen that short after these peaks some valleys with low occupation appear. The idea of delaying some connections and processing them in a future slot is the central idea of this work.
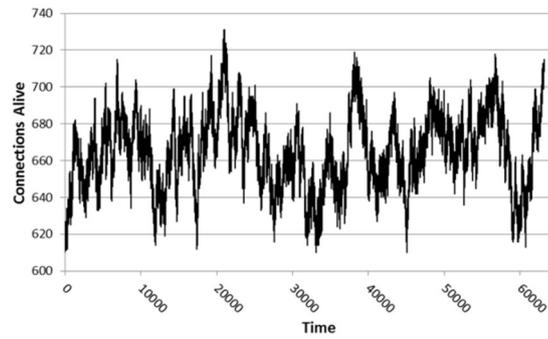

Figure 1. Instantaneous number of living connections during simulation time in a typical EON (NSF topology, BBP=1%, uniformly distributed connections with bandwidth ranging from 1 to 8 spectrum slots).

Some previous works have dealt recently with time constraints in EONs [2, 7, 8]. The main concept supporting all of these works is Advanced Reservation (AR), which has been widely used in resources allocation in other network problems since long time ago [9]. AR assumes that connections arrival is known in advance and reserves the required resources with some anticipation. Therefore, connections scheduling becomes an additional constraint to the RSA algorithm. Authors in [2] propose an algorithm to optimize resources usage when the generated connections have rigid start and ending times or can be flexibly allocated during a given time window. Of course later traffic gives more freedom and chances to find optimum resources allocation distribution. On the other hand, traffic in dynamic scenarios is generally unpredictable as the arriving instant, duration and capacity of future requests are stochastically distributed and not known in advance. Different RSA strategies to improve EON performance by exploiting the holding time awareness are demonstrated in [7]. Immediate reservation (IR) requests are mixed with advanced reservation (AR) requests and some ideas such as pre-emption (IR traffic is dropped when AR asks for its assigned resources) are introduced and evaluated. Differences in WDM network performance are observed and reported when changing the book-ahead time of reservations. A complete RMSA and scheduling in EONs supporting multi-class (IR and AR) traffic is introduced in [8]. Different methods and strategies in the allocation of these multi-class connections are evaluated and network performance trends considering not only blocking but also delay are given.

Our approach is slightly different from previous works in that we do not apply advanced reservation, but our focus is on evaluating the influence of delay tolerant traffic (DTT) on the blocking performance of the EON. Indeed, a given fraction (ranging from 0 to 100%) of the generated connections are considered to be DTT. Nevertheless, these connections are not scheduled for the future but, in case of not finding free resources at their birth instant, they are just buffered and their allocation is retried after a certain time period. It is shown by simulation how the different parameters involved affect the network performance, and some ideas about required resources in a realistic scenario are provided. Effects of parameters such as waiting time to retry or buffering capacity needed at the network nodes are analysed.

3. SIMULATION FRAMEWORK AND RESULTS

In order to evaluate the impact that the presence of DTT has on the EONs performance, an ad-hoc Matlab® based simulator has been deployed. Its main characteristics are,

- Number of FS per connection (BW) ranging from 1 to 8. Concerning the connections BW statistics Uniform and Normal distribution of connections have been tested. Each generated connection is randomly labelled as delay-tolerant or not. The probability of connections being marked as delay tolerant (DTT) is adjusted and it is the key parameter under study. Depending on this probability each new generated connection can either be delay tolerant or not. If at the RSA stage there are no free resources to allocate it, it will be blocked or, in case it has been marked as DTT, stored for a second attempt after some waiting time.

- NSFnet (14 nodes, 21 bi-directional links), European Optical Network (EONet, 19 nodes, 37 bi-directional links) and USBackbone (24 nodes, 48 bi-directional links) topologies [4] have been simulated. Connections Source-Destination pairs are uniformly distributed among all the network's nodes.

- Regarding the routing and spectrum allocation (SA) algorithm for allocating new connections, simulations run a typical adaptive shortest path routing and First-Fit SA. In case of blocking at the first path attempted the most crowded link is temporarily withdrawn and up to two new path searches are done.

- Number of FS (each one corresponding to 12.5 GHz) per link is equal to 320, so a 4 THz spectrum band is used in each fiber. This value corresponds approximately to the optical fibers C-band.

- The connection's inter-arrival time (IAT) follows a Poisson process with an average value of IAT=1 t.u. (time unit). The connection's holding time follows a negative exponential distribution. The average value (HT t.u.) is adjusted to obtain the appropriate offered load values (which in turn change the BBP).

As previously explained, the main objective of this work consists in evaluating the effects of DTT as well as how these effects depend on some parameters such as waiting time for retrial or buffering capacity at the network nodes. First results are shown in Fig. 2. In Fig. 2.(a) all the traffic has been considered to be DTT and the effect of the waiting time (before retrying allocation) on the BBP has been tested for the three different topologies simulated. It has been found a similar behaviour regardless of the considered topology. Blocking rapidly reduces when waiting time starts growing (waiting time equal to zero is equivalent to having DTT=0) and only a marginal reduction is observed when waiting time is greater than half of the average connections holding time (0.5*HT). Taking this result into account, the waiting time is fixed to 0.6*HT during next simulations. On the other hand, these simulations will only consider NSF topology, as it has been demonstrated similar performance on the different topologies tested. The gain in terms of supported load at BBP=1% is between ranges from 11% to 13% depending on the topological characteristics of each network (EONet performs slightly worse).
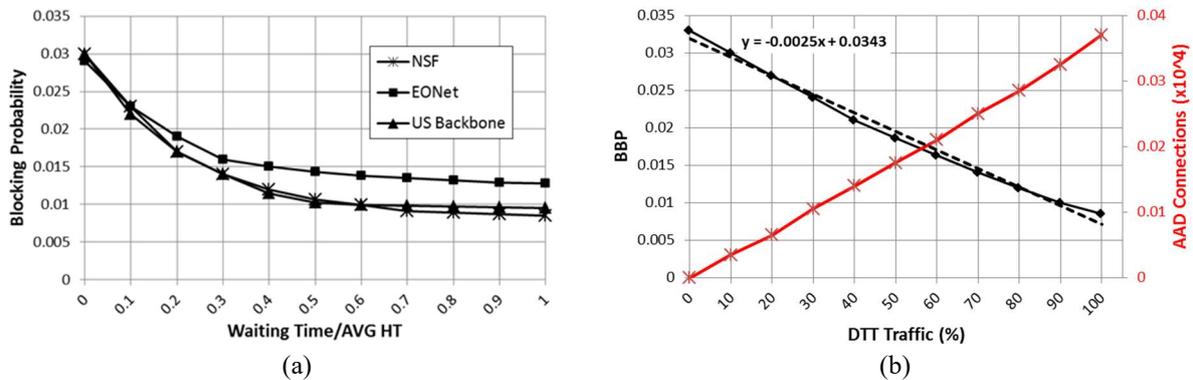

(a) (b)

Figure 2. (a) Effect of DTT on the BBP vs. waiting time before allocation retrial. Three different networks have been simulated. (b) Bandwidth blocking probability reduction (left axis) and number of connections allocated after delay (right axis) vs. percentage of DTT traffic.

Bandwidth blocking probability (BBP) reduction is shown as a function of the percentage of DTT traffic in Fig. 2.(b). As stated before, waiting time in this case is set to 0.6*HT. Main conclusion from Fig. 2.(b) is the linear dependency of BBP with DTT percentage. It is also shown on the right axis the total number of connections that have been buffered and re-allocated during the simulation (the total number of connections generated in each of these simulations was 20000). It can be seen on the one hand how BBP is reduced from 3.3% (when no DTT is present) to less than 1% (when all the connections are DTT), showing an almost perfect linear behaviour. On the other hand, the number of connections which are allocated after delay (AAD) reaches a value close to 400 (around a 2% of the total number of generated connections) when 100% of the connections are DTT.

Fig. 3.(a) focuses on the effect of nodes buffering capacity. The number of transmitters per node has been fixed to ensure that no blocking is due to lack of transmitters (it has been found that having 50 transponders per node

guarantees it). On the other hand, it is shown that having a buffering capacity of 24 FS per node ensures that the blocking due to buffers limitation is almost negligible.
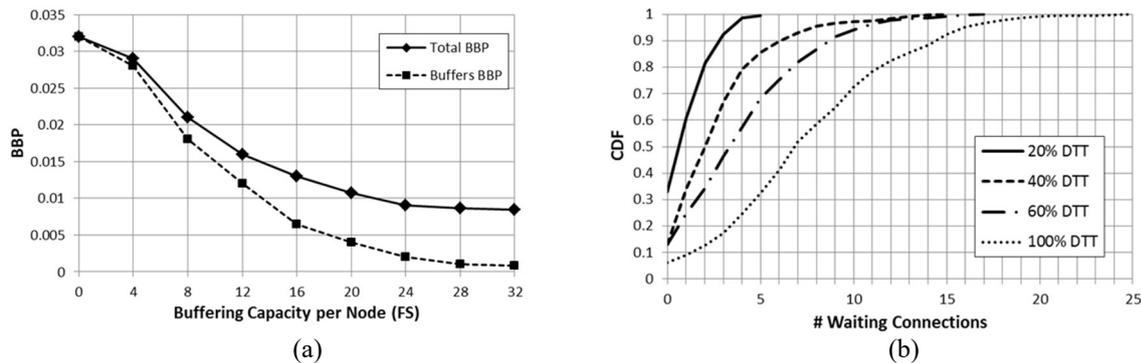

(a)   (b)

Figure 3. (a) Effect of nodes buffers size on the blocking probability. (b)Cumulative distribution function of the instantaneous number of connections waiting during simulation.

Last simulations (shown in Fig. 3.(b)) were carried out with the aim of measuring the number of connections waiting for delayed allocation during simulations. When DTT=40% a peak of 15 connections waiting was found, even the average number of waiting connections during the whole simulation was as low as about 3. This peak reached 25 connections when DTT=100%. The average number of waiting connections in this case was almost 8. From all of the obtained results it can be inferred that having DTT traffic in EONs reduces considerably the BBP. This reduction translates to an increase of the load supported at a given BBP. Adapting the waiting time to the instantaneous network status and introducing any de-fragmentation mechanism [5] in the network are left for future study.

4. CONCLUSIONS

Real world traffic is composed by some connections with strict delay requirements and others which can be delayed without severe penalties. Knowing that traffic in EONs shows a very variant behaviour (randomness in connections birth and death implies huge variations in the network load), this work has studied the impact of delay tolerance of traffic in EONs performance. It has been shown that, under generic dynamic conditions, a reduction of up to 75 % (from 3.3% to 0.008%) is achieved in terms of BBP. This is equivalent to an increase of the load supported higher than 10%.

REFERENCES

[1]   O. Gerstel, M. Jinno, A. Lord, B. Yoo, "Elastic Optical Networking: A New Dawn for the Optical Layer," IEEE Commun. Mag., vol. 50, no. 2, pp. s12-s20, 2012.
[2]   P. Afsharlar, A. Deylamsalehi, J. M. Plante, J. Zhao, V. M. Vokkarane, "Routing and Spectrum Assignment With Delayed Allocation in Elastic Optical Networks," Journal of Optical Communications and Networking, Vol. 9, No. 3, pp. B101-B111, 2017.
[3]   B. C. Chatterjee, N. Sarma, E. Oki, "Routing and Spectrum Allocation in Elastic Optical Networks: A Tutorial, " IEEE Communications Surveys & Tutorials, Vol. 17, pp. 1776-1800, 2015.
[4]   J. Comellas, L. Vicario, G. Junyent, "Using Static Connections to Improve Elastic Optical Networks Performance," OSA Advanced Photonics Congress, paper NeTu1B.5, 2017.
[5]   J. Comellas, L. Vicario, G. Junyent, "Proactive Defragmentation in Elastic Optical Networks under Dynamic Load Conditions", Photonic Network Communications, Vol. 36, pp. 26-34, 2018.
[6]   K. Christodoulopoulos, I. Tomkos, E. Varvarigos, "Dynamic Bandwidth Allocation in Flexible OFDM-based Optical Networks," in Proc. OFC, 2011, OTuI5.
[7]   A. Muhammad, R. Forchheimer, "Coexistence of Advance and Immediate Reservation in WDM Networks: Some RWA Strategies", in Proc. ICTON 2012, Tu.C2.4.
[8]   E. E. Moghaddam, H. Beyranvand, J. A. Salehi, "Routing, Spectrum and Modulation Level Assignment, and Scheduling in Survivable Elastic Optical Networks Supporting Multi-Class Traffic", Journal of Lightwave Technology, Vol. 36, pp. 5451-5461, 2018.
[9]   E. Escalona, S. Spadaro, J. Comellas, G. Junyent, "Advance reservations for service-aware GMPLS-based optical networks", Computer Networks, Vol. 52, pp. 1938–1950, 2008.