Application of Tell & Go and Tell & Wait Reservation Strategies in a Optical Burst Switching Network: a Performance Comparison

A. Detti*, M. Listanti°

°University of Rome "La Sapienza", INFOCOM Dept., Italy

Tel: +39 06 44585458; e-mail: marco@infocom.ing.uniroma1.it

* University of Rome "Tor Vergata", Dip. Ing. Elettronica, Italy

ABSTRACT

This paper deals with the comparison between two resource reservation techniques, named Tell and Go (TG) and Tell and Wait (TW), when they are used in an All Optical Network for the support of high-speed IP traffic. This comparison study is motivated considering that the modes currently under investigation for the implementation of the next generation coarse packet switched optical networks, e.g. Burst Switching, are based on these reservation policies. The comparison methodology here proposed takes into account basic parameters tied to the available optical technology and is independent of the traffic model. The figure of merit which the comparison is based on is the traffic gain defined as the ratio between the amount of offered traffic supported by the two reservation techniques under the same bottleneck link blocking probability. The obtained results provide a practical guideline for the identification of the relevant applicability scenarios of the two techniques. In particular, it shows that with the current state of art of the optical switch devices, the potential advantages of TG technique only arise in wide are networks.

I. INTRODUCTION

In these last years some researches have been focussed on the definition of a new optical network paradigm to be used in the future *All Optical Network* (AOL) [1] that be able to cope with the "next wave" of IP traffic. It's foreseeable that, in the first phase the AON will offer high capacity circuit switched services by means of the provisioning of WDM end-to-end optical paths [2]. Unfortunately, such a solution is unsuitable for highly bursty traffic like the IP one because its intrinsic inefficiency due to the coarse granularity of a wavelength bandwidth. A great research effort has been focused on the optical packet switching operating with fixed length packets; nevertheless, the currently unavailability of high capacity optical storage devices, makes this technique difficult to be implemented in a short perspective. In a nearer term scenario, bufferless asynchronous switching techniques based on variable length containers, called *bursts*, seem to be more promising. An example of these asynchronous techniques is the *Optical Burst Switching* (OBS) [3,4,5,6,7].

Although, the OBS has been extensively studied in several aspects, i.e. node architectures, burst scheduling policies, etc., the definition of the most suitable resource reservation strategy is still an open issue. This paper aims at giving a contribution in this direction by comparing two possible resource reservation strategies, namely *Tell and Wait* (TW) and *Tell and Go* (TG), candidated to be used in a *Burst Switched Optical Network* (BSON).

Previous works [8,9] showed that in an high speed environment, the TG strategy is more efficient than the TW one due to the short transmission times in comparison with the end-to-end propagation delays. This conclusion is achieved by hypothesizing electronic switching systems in which the times required to configure the switching fabric of a node are negligible with respect to the transmission and propagation times. In an optical environment, the aforementioned hypothesis does not seem to hold anymore; in fact, at the current state of the art, the configuration time of an optical device can be even of the order of a few milliseconds [13]. The aim of the paper is to dip the comparison of the two techniques in a specific technological scenario and to determine the relevant convenience regions taking into account: i) the constraints imposed by the optical technology; ii) the network dimensions and iii) the burst sizes.

Moreover, it is to be noted that the TG can lead to implementation difficulties, as a matter of example, in [9] it is proved that the TG is potentially unstable if the offered traffic is not limited and then specific edge functions are needed. Therefore, we are interesting to find the technological and network conditions for which the supposed higher complexity of the TG strategy is justified by a considerable improvement of the transport capability.

In section II the reference optical network framework is outlined. In section II.A and II.B, the TW and TG techniques are briefly summarized. In section III the comparison methodology is presented, whereas in sections IV.B and IV.C, the model is applied in the cases of the ring topology and of the *vBNS* backbone network. Finally conclusion are given in section V.

II. OPTICAL BURST SWITCHING NETWORK MODEL

The considered network model is shown in Fig. 1. It consists of a backbone network operating according to the *Optical Burst Switching* (OBS) paradigm and offering a transparent transport service to the adjacent electronic IP networks. The interface functions between the electrical and optical worlds are performed by the edge nodes (*ENs*).

The functions of an ingress *EN* are: i) to build the optical bursts by aggregating a number of IP packets directed to

the same egress *EN*; ii) to insert within the bursts the suitable information for delineating and extracting the IP packets [12] at the receiving node; iii) to reserve the needed resources for the burst transfer through the optical backbone; iv) to forward the bursts through the network.
Vice versa, an egress *EN* delineates and recovers the IP packets contained in the received bursts and forwards them towards the destination networks.

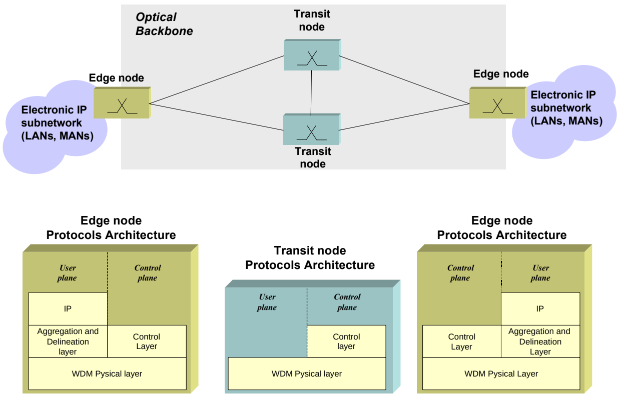

Fig. 1: Optical Network Framework

A link of the optical backbone is a fiber supporting a WDM multiplex with *W+n* wavelengths: *W* wavelengths, called *data wavelengths*, are dedicated to the burst transmission, whereas the remaining *n*, called *control wavelengths*, are signaling channels devoted to the transport of the control messages between the network nodes. The control messages contain the information allowing a node to route the bursts and to properly configure the switching fabric of the crossed nodes.
A transit node logically consists of a bufferless optical switching fabric and an electronic centralized control unit. The switching fabric is equipped with wavelength converters in order to solve the output contentions between incoming bursts. The control unit electronically processes the control messages and performs the related actions according to the adopted resource reservation strategy. Two reservation strategies are here considered: the Tell and Wait (TW) and the Tell and Go (TG). These two strategies are briefly summarized in the following.

II.A. Tell and Wait resource reservation strategy

In the Tell and Wait (TW) technique, the a burst is emitted by an ingress node only if an optical virtual path has been set through the network up to the egress node. An optical virtual path is defined by the concatenation of the wavelengths reserved, link by link, during the set up phase preceding the burst transmission. In this sense, the TW technique corresponds to the mode of operation of the traditional circuit switching.
As shown in Fig. 2, when an ingress *EN* has a new burst to transfer, it sends, on a control wavelength, a *setup* control message towards the egress *EN*. This message aims at reserving a data wavelength on each link along the path between the ingress and the egress node. When the *setup message* is received by a transit node, the control unit reserves a free data wavelength on the routed output; such a data wavelength is dedicated to the burst as long as an explicit *release message* is received. Once the reservation has been performed, the control unit configures the node switching fabric possibly determining the proper wavelength conversion to solve the output contentions with other incoming bursts.

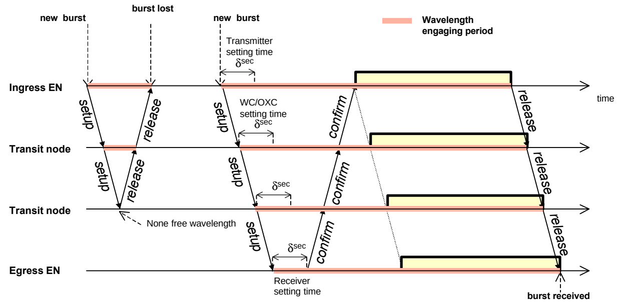

Fig. 2: Tell and Wait

If a free data wavelength is found on each link, a *confirm message* is emitted by the egress node in the reverse direction with the aim at notifying the ingress *EN* of the success of the setup phase of the optical virtual path. Each intermediate node along the path will forward the *confirm message* only if the configuration of the optical switching fabric is terminated; otherwise, the forwarding of the *confirm message* is delayed.
When the ingress *EN* receives the *confirm message*, it transmits the burst on the previously reserved optical virtual path; as soon as the burst transmission is over, the node emits a *release message*. This message aims at tearing down the optical virtual path and at freeing the wavelengths reserved along the path. Therefore, a wavelength on a link is engaged for a time period bounded between the receptions of the *setup* and the *release messages*.
If the setup phase can not successfully terminated, i.e. no free data wavelength is found on a link, a *release message* is immediately emitted in the reverse direction, towards the ingress node. This message immediately releases the previously engaged wavelengths and notifies the ingress *EN* of the reservation unsuccessfulness, that consequently discard the burst. The lost burst may be recovered or by the higher layer protocols (e.g. TCP) or by some kind of optical channel layer protocols [15].

II.B. Tell and Go resource reservation strategy

In the Tell and Go (TG) solution, a burst is emitted even if the establishment of an optical virtual path has not been completed. The burst follows the virtual path while the setup phase is in progress using those wavelengths that have been already engaged. If the setup phase will be successfully completed the burst will arrive at destination, otherwise it will be discarded in an intermediate node.
As shown in *Fig. 3*, when an ingress *EN* has a new burst to transfer, sequentially sends towards the egress *EN* the *setup message*, the burst and the *release message*. The setup message and the burst will be spaced by a guard time at least equal to the time interval needed for the configuration of the optical switching fabric inside a node. This guard time allows the optical devices to be set before the burst arrival. The meaning of the *setup* and *release* messages and the relevant actions are the same as

in the TW case. If a free wavelength is not found on a link along the ingress-egress path, the burst is discarded by the node preceding the blocked link. No *release message* is emitted towards the ingress *EN*.

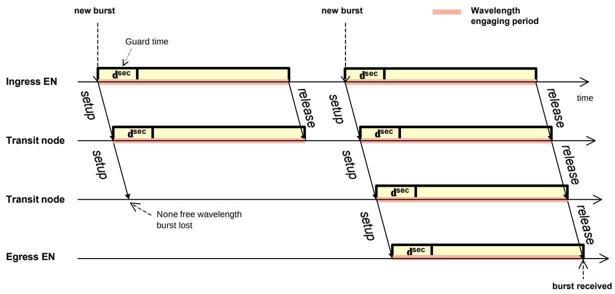

Fig. 3: Tell and Go

The TG technique is a simplified version of the *Just-Enough-Time* (JET) [4,5,6] proposed for the Optical Burst Switching. In JET an additional *offset time* is inserted between the *setup message* and the burst; this time takes into account the processing and transmission delays experienced by the setup message within a node. Such a time is needed if a node is not equipped with fiber delay lines aiming at delaying the incoming bursts of a time interval equal to that suffered by the relevant setup message. In our model, as it will be clarified later, since the processing and transmission delays of the control messages are neglected, the TG scheme is equivalent to the JET.

III. COMPARISON METHODOLOGY AND ANALYTICAL MODEL

The goal of this section is to describe the methodology and the analytical model adopted for the comparison of the performance of the TW and TG techniques. The comparison aims at determining, with a specific reference to an optical network scenario, the most suitable areas for the application of the two techniques taking into account the constraints imposed by the current technology, the network dimensions and the burst sizes. The performance parameter utilized for the comparison is the amount of the traffic, in terms of bursts per unit time, that can be carried by the backbone network adopting the TW or the TG techniques with the constraint to be subjected to the same burst loss probability. In particular, it is imposed that the loss probability on the *bottleneck link* is equal for the two techniques. The bottleneck link is defined as network link characterized by the highest value of the loss probability.

The comparison is carried out under the following conditions:

- c1) all the network links have the same number of data wavelengths;
- c2) the bit rate is equal on all the wavelengths of every link;
- c3) the configuration time of the optical devices is the same for all of the nodes, it is indicated as d^{sec} and is supposed equal to the guard time between the *setup message* and the burst in the TG strategy;
- c4) the processing and transmission times of a control message are supposed to be negligible, i.e. the processing power of the control unit of a node is considered to be infinite;
- c5) the traffic space distribution, i.e. the probability that a burst belongs to a specific traffic relation, defined by the couple of ingress and egress ENs, is the same for the two techniques;
- c6) the routing of the bursts belonging to a specific traffic relation, i.e. the sequence of the crossed optical links, is fixed and equal for both the techniques; in the rest of the paper the route followed by the bursts belonging to a traffic relation will be indicated as *optical path*, or simply *path*, and the same term will be also used also to indicate the traffic relation itself;
- c7) the burst loss probability l_{worst} on the bottleneck link is equal for both the techniques.

From the performance point of view, the key difference between the two techniques consists in the different amount of network resources consumed for a burst transmission.

As far as the TW technique is concerned, as results from Fig. 2, in case of a successful setup, a wavelength on each link of the path is occupied for the transmission of a burst for a time interval equal by the sum of three terms: a) the burst transmission time; b) the optical devices configuration time d^{sec}; c) the end-to-end round trip time. In case of unsuccessful setup, the TW technique involves the useless occupation of a wavelength on every link preceding the blocked one for a time interval equal to the round trip time between a link and the blocked one.

As for the TG technique (see *Fig. 3*), the resource consumption on a link of the path in case of a successful burst transmission is given by the sum of the guard time d^{sec} and the burst transmission time. Instead, if the virtual path can not be found, the TG wastes a wavelength for the same amount of time on every link preceding the blocked one. It is to be noted that, in case of TG technique, the end-to-end round trip time has no influence on the resource consumption since the burst is emitted while the path setup phase is still in progress.

It is to be noted that, in the following, the term *offered traffic* will be used to indicate the mean amount of network resources required for the burst transmission without considering any protocol overhead; whereas, the term *offered load* will indicate the mean amount of resources needed to support the offered traffic according to the chosen protocol. In other words, the offered load is the offered traffic plus the protocol overheads (round trip time, guard times, etc.).

The analytical model is based on the following assumptions:

- h1) the network is in the steady state condition;
- h2) link independence, i.e. the status of a link is independent of that of the others;
- h3) the load that a path offers to a link is thinned by the loss occurred in the other links of the path [9];
- h4) if the amounts of load that the TG and TW techniques offer to a network link are equal, then the related link loss probabilities are the same.

Let us introduce some notations:

- M : number of network links; each link is uniquely identified by an index j ($1 \le j \le M$);
- W: number of per-link data wavelengths;
- T_{burst} : mean value of the duration of the bursts;
- P : number of *ingress/egress EN* network paths (or traffic relations), each path is uniquely identified by an index i ($1 \le i \le P$);
- $H=\{h_i\}$: *path length vector*; h_i ($1 \le i \le P$) is the length of the path #i measured in number of hops;
- S: the number of hops of the longest network path; that is $S = \max_i \{h_i\}$;
- $B=\{b_{i,j}\}$: (PxM) *path-link incidence matrix*; $b_{i,j}=1$ if the link #j belongs to path #i, otherwise $b_{i,j}=0$;
- $Z=\{z_{i,j}\}$: (PxM) *path-link position matrix*, if $b_{i,j}=1$, $z_{i,j}$ ($1 \le z_{i,j} \le h_i$) is the position of the link #j along the path #i, otherwise $z_{i,j}=0$;
- $G=\{g_{i,x}\}$: (PxS) *path-link sequence matrix*; if $x \le h_i$, $g_{i,x}$ ($1 \le g_{i,x} \le M$) indicates the index of the link placed in position #x along the path #i; otherwise $g_{i,x}=0$;
- $R=\{r_j\}$: *normalized link propagation delay vector*; r_j ($1 \le j \le M$) is the propagation delay over the link #j normalized to T_{burst};
- $A^X = \{a_i^X\}$: *path offered traffic vector*; a_i^X (X=TW, TG; $1 \le i \le P$) is the amount of external traffic, measured in Erlangs, offered over the path #i in case of TW or TG technique, i.e. the number of offered bursts in a time interval equal to T_{burst};
- $a_{Tot}^X = \sum_{i=1}^{P} a_i^X$: *total network offered traffic* (X=TW, TG), the total amount of external traffic, measured in Erlangs, offered to the optical backbone in case of TW or TG technique;
- $A^s = \{a_i^s\}$: *statistical path offered traffic vector*; a_i^s ($1 \le i \le P$) is the probability that an offered burst belonging to the path #i; on the basis of conditions c5 this vector is identical for both the techniques, hence $A^X = a_{Tot}^X \cdot A^S$ (X=TW, TG);
- $C^X = \{c_{i,j}^X\}$: *path-link load matrix*; $c_{i,j}^X$ (X=TW, TG; $1 \le i \le P$; $1 \le j \le M$) is the load, measured in Erlangs, offered from the path #i to the link #j in the TG or TW cases;
- $L^X = \{l_j^X\}$: *link loss probability vector*; l_j^X (X=TW, TG; $1 \le j \le M$) is the probability that a burst offered to the link #j is loss in the TG or TW cases;
- $CL^X = \{CL_j^X\}$: *link load vector*; CL_j^X (X=TW, TG; $1 \le j \le M$) is the total load over the link #j measured in Erlangs in the TG or TW cases;
- $CL_{worst}^X = \max_j [CL_j^X]$: (X=TW, TG) the load over the bottleneck link for the TW and TG techniques.
- d: the optical device configuration time normalized to T_{burst}, i.e. $d = d^{\sec} / T_{burst}$.

As above mentioned, the analytical model has the goal to evaluate the traffic gain T_G defined as the ratio between the amount of offered traffic that can be carried by the network in case of TG and TW techniques, with the constraint to obtain the same burst loss probability l_{worst} on the bottleneck link. On the basis of the previous definition we have:

$$T_G = \frac{a_{Tot}^{TG}}{a_{Tot}^{TW}} \tag{1}$$

Let's assume that, for the TW and TG techniques, the bottleneck links are the number #k and the number #h, respectively; so we have:

$$CL_{worst}^{TW} = CL_k^{TW} ; \qquad CL_{worst}^{TG} = CL_h^{TG} \; . \tag{2}$$

According to the (c7) condition, the blocking probabilities on both the bottleneck links must be the same, then, on the basis of assumption (h4), the loads on these links must be the same. Hence

$$CL_{worst}^{TW} = CL_{worst}^{TG} \tag{3}$$

In sections III.A and III.B, the values of CL_{worst}^{TW} and CL_{worst}^{TG} are determined as function of the couples a_{Tot}^{TW}, L^{TW} and a_{Tot}^{TG} and L^{TG}, respectively, moreover, the expressions of the upper and a lower bounds of both quantities are found, finally, in section III.C, an upper and a lower bound of T_G making use of the previously mentioned bounds are determined.

III.A. Evaluation of the bottleneck link load in case of the TW technique

In this section, we express CL_{worst}^{TW} as a function of the a_{Tot}^{TW} and L^{TW}. Let P_k be the set of paths crossing the bottleneck link #k. The load CL_k^{TW} is given by the sum of the loads offered to this link from all the paths belonging to P_k, i.e.

$$CL_k^{TW} = \sum_{i \in P_k} c_{i,k}^{TW} \tag{4}$$

Adopting the same line of reasoning of [9], the load $c_{i,k}^{TW}$ can be expressed as follows:

$$\begin{aligned} c_{i,k}^{TW} &= a_i^{TW} \cdot \sum_{s=z_{i,k}+1}^{h_i} \left[\prod_{\substack{x=1 \\ x \ne z_{i,j}}}^{s-1} \left(1 - l_{g_{i,x}}^{TW}\right) \cdot l_{g_{i,s}}^{TW} \cdot 2 \cdot \sum_{x=z_{i,k}}^{s-1} r_{g_{i,x}} \right] + \\ &+ a_i^{TW} \cdot \left[\prod_{\substack{x=1 \\ x \ne z_{i,k}}}^{h_i} \left(1 - l_{g_{i,x}}^{TW}\right) \cdot \left(1 + d + 2 \cdot \sum_{x=1}^{h_i} r_{g_{i,x}} \right) \right] \end{aligned} \tag{5}$$

Let's define the *normalized path-link load* $\tilde{c}_{i,k}^{TW}$ and the *normalized bottleneck link load* $\tilde{C}_{worst}^{TW}$ dividing (5) and (4) by a_{Tot}^{TW}, so we have:

$$\tilde{c}_{i,k}^{TW} = a_i^s \cdot \sum_{s=z_{i,k}+1}^{h_i} \left[\prod_{\substack{x=1 \\ x\neq z_{i,j}}}^{s-1} \left(1-l_{g_{i,x}}^{TW}\right) \cdot l_{g_{i,s}}^{TW} \cdot 2 \cdot \sum_{x=z_{i,k}}^{s-1} r_{g_{i,x}} \right] + \quad (6)$$

$$a_i^s \cdot \left[\prod_{\substack{x=1 \\ x\neq z_{i,k}}}^{h_i} \left(1-l_{g_{i,x}}^{TW}\right) \cdot \left(1+d+2\cdot\sum_{x=1}^{h_i} r_{g_{i,x}}\right)\right]$$

$$\tilde{C}L_k^{TW} = \sum_{i\in P_k} \tilde{c}_{i,k}^{TW} \quad (7)$$

Therefore the TW bottleneck link CL_{worst}^{TW} load can be rewritten as:

$$CL_{worst}^{TW} = a_{Tot}^{TW} \cdot \tilde{C}L_{worst}^{TW} = a_{Tot}^{TW} \cdot \max_j [\tilde{C}L_j^{TW}] \quad (9)$$

It is to be noted that, $\tilde{C}L_{worst}^{TW}$ can be thought as the load of the bottleneck link for an unitary network offered traffic when the links loss probabilities are set equal to those obtained for a network offered traffic a_{Tot}^{TW} (those given by the vector L^{TW}). Obviously, this doesn't mean that for an unitary network offered traffic the $\tilde{C}L_{worst}^{TW}$ is the value of the bottleneck link load because the links loss probabilities may be different from those indicated by the vector L^{TW}.

The exact evaluation of C_{worst}^{TW} requires the application of a long and tedious iterative numerical procedure that does not provide a tractable analytical expression, so, it is useful to find an upper and a lower bound of C_{worst}^{TW}.

Let #*u* and #*v* be two generic links of the network. It is straightforward to prove that, the load on the link #*u* offered by a path #*i* crossing that link, is a not-increasing function of the loss probability of the link #*v*. As a matter of fact, the derivative of the load offered from the path #*i* to the link #*u*, with respect to the blocking probability on the link #*v*, is less or equal to zero; in formula $\partial c_{i,u}^{TW} / \partial l_n^{TW} \le 0$. As the link load is the sum of the loads offered by the various crossing paths, then it is possible to affirm that the load on link #*u* is a no-increasing function of the loss probability of link #*v*; in formula $\partial CL_u^{TW} / \partial l_n^{TW} \le 0$.

On the basis of the previous considerations, fixed the total network offered traffic a_{Tot}^{TW}, a lower and an upper bounds of the load on the bottleneck link can be found as follows:

1. <u>Lower bound</u>: if the number of wavelengths of the network links is modified in order to obtain a loss probability equal to l_{worst} in every network links, all the link loss probabilities increase or, at least, do not decrease; therefore, the loads on the various network links will be less or equal to the actual ones given by CL^{TW}; mathematically, if $CL^{TW}(l_i^{TW} = l_{worst}, 1 \le i \le M)$ is the link load vector computed by means of (4) supposed that each link loss probability is fixed equal to l_{worst}, then it results $CL_j^{TW}(l_i^{TW} = l_{worst}, 1 \le i \le M) \le CL_j^{TW}$ (1≤*j*≤*M*), as a consequence, the lower bound of the load on the bottleneck link is given by:

$$CL_{worst}^{TW}(low) = \max_j [CL_j^{TW}(l_i^{TW} = l_{worst}, 1 \le i \le M)] \le CL_{worst}^{TW} \quad (10)$$

2. <u>Upper bound</u>: if the number of wavelengths of the network links is modified in order to obtain a loss probability equal to zero in every network links, all the link loss probabilities decrease or, at least, do not increase; therefore now, the loads on the various network links will be higher or equal to the actual ones given by CL^{TW}, mathematically, if we call $CL^{TW}(l_i^{TW} = 0, 1 \le i \le M)$ the link load vector computed by the (4) with each link blocking probability equal to zero, then it results $CL_j^{TW}(l_i^{TW} = 0, 1 \le i \le M) \ge CL_j^{TW}$ (1≤*j*≤*M*), as a consequence, the upper bound of the load on the bottleneck link is given by:

$$CL_{worst}^{TW}(up) = \max_j [CL_j^{TW}(l_i^{TW} = 0, 1 \le i \le M)] \ge CL_{worst}^{TW} \quad (11)$$

Therefore, on the basis of (10) and (11), the bottleneck link load CL_{worst}^{TW} is bounded by:

$$CL_{worst}^{TW}(low) \le CL_{worst}^{TW} \le CL_{worst}^{TW}(up) \quad (12)$$

Taking into account (9), from (12), we obtain the upper $\tilde{C}L_{worst}^{TW}(up)$ and lower $\tilde{C}L_{worst}^{TW}(low)$ bounds of the normalized bottleneck link load $\tilde{C}_{worst}^{TW}$:

$$\tilde{C}L_{worst}^{TW}(low) \le \tilde{C}L_{worst}^{TW} \le \tilde{C}L_{worst}^{TW}(up) \quad (13)$$

wherein

$$\tilde{C}L_{worst}^{TW}(low) = \tilde{C}L_{worst}^{TW}(l_i^{TW} = l_{worst}, 1 \le i \le M) \quad (13a)$$

$$\tilde{C}L_{worst}^{TW}(up) = \tilde{C}L_{worst}^{TW}(l_i^{TW} = 0, 1 \le i \le M) \ . \quad (13b)$$

III.B. Evaluation of the bottleneck link load in case of the TG technique

Applying the same procedure as in the TW case and remembering that the link #*h* represents the bottleneck link, it follows

$$CL_h^{TG} = \sum_{i\in P_h} c_{i,h}^{TG} = [a_i \cdot (1+d)] \cdot \prod_{x=1}^{z_{i,h}-1} (1 - l_{g_{i,x}}^{TG}) \quad (14)$$

The normalized values are given by:

$$\tilde{c}_{i,h}^{TG} = \left[a_i^s \cdot (1+d)\right] \cdot \prod_{x=1}^{z_{i,h}-1} (1 - l_{g_{i,x}}^{TG}) \quad (15)$$

$$\tilde{C}L_j^{TG} = \sum_{i\in P_h} \tilde{c}_{i,h}^{TG} \quad (16)$$

Therefore the TG bottleneck link load CL_{worst}^{TG} can be rewritten as :

$$CL_{worst}^{TG} = a_{Tot}^{TG} \cdot \tilde{C}L_{worst}^{TG} = a_{Tot}^{TG} \cdot \max_j [\tilde{C}L_j^{TG}] \quad (17)$$

As in the TW case, it is useful to reply the same reasoning in order to obtain the upper and lower bounds of the bottleneck link load. The lower bound $CL_{worst}^{TG}(low)$ can be evaluated from the (17) by imposing that all the links loss probabilities equals l_{worst}, whereas the upper bound $CL_{worst}^{TG}(up)$ can be evaluated with the (17) by assuming that all the links blocking probabilities equals zero. This obviously is true even for the normalized values $\tilde{C}L_{worst}^{TG}(low)$ and $\tilde{C}L_{worst}^{TG}(up)$. So, we have :

$$CL_{worst}^{TG}(low) \le CL_{worst}^{TG} \le CL_{worst}^{TG}(up) \quad (18)$$

wherein,

$$CL_{worst}^{TG}(low) = \max_j [CL_j^{TG}(l_i^{TG} = l_{worst}, 1 \le i \le M)] \quad (18a)$$

$$CL_{worst}^{TG}(up) = \max_j [CL_j^{TG}(l_i^{TG} = 0, 1 \le i \le M)] \quad (18b)$$

and

$$\tilde{C}L_{worst}^{TG}(low) \le \tilde{C}L_{worst}^{TG} \le \tilde{C}L_{worst}^{TG}(up) \quad (19)$$

wherein

$$\tilde{C}L_{worst}^{TG}(low) = \tilde{C}L_{worst}^{TG}(l_i^{TG} = l_{worst}, 1 \le i \le M) \quad (19a)$$

$$\tilde{C}L_{worst}^{TG}(up) = \tilde{C}L_{worst}^{TG}(l_i^{TG} = 0, 1 \le i \le M) \ . \quad (19b)$$

III.C. *Traffic gain evaluation*

Now we have all the expressions to evaluate the traffic gain T_G, as a matter of fact, utilizing the (9) and (17), we are able to rewrite (3) as follows:

$$\frac{\tilde{C}L_{worst}^{TW}}{\tilde{C}L_{worst}^{TG}} = \frac{a_{Tot}^{TG}}{a_{Tot}^{TW}} \quad (20)$$

So, the traffic gain can be expressed as:

$$T_G = \frac{\tilde{C}L_{worst}^{TW}}{\tilde{C}L_{worst}^{TG}} \quad (21)$$

As above explained, we are not interested to evaluate the actual value of T_G, but we can easily bound its value by means of (13) and (19), so

$$T_G(low) \le T_G \le T_G(up) \quad (22)$$

wherein

$$T_G(low) = \frac{\tilde{C}L_{worst}^{TW}(low)}{\tilde{C}L_{worst}^{TG}(up)}, \quad T_G(up) = \frac{\tilde{C}L_{worst}^{TW}(up)}{\tilde{C}L_{worst}^{TG}(low)}, \quad (23)$$

In section IV.C, we show that the tightness of the two bounds are quite good, in fact if l_{worst} assumes low values, let's say lower than 10^{-3}, the values of the two bounds are undistinguishable.

IV. COMPARISON RESULTS : NETWORK SIZE ANALYSIS

In this section the results of the comparison between the TW and TG techniques is carried out with the aim at highlighting the impact of network dimensions. i.e. the geographical area covered by the network. In order to provide the results in a compact form, a generic network in which all the links have the same length is considered; obviously, in such a network, the propagation delays over all of the links are equal. In the following we express the length of the links by means of the relevant propagation delay. Let r and δ be the normalized values of the propagation delay over the network links and of the configuration time of the optical devices in each node, both the values are normalized with respect to T_{burst}.

The comparison is based on the analysis of the iso-gain curves $r(d, b)$, i.e. the set of points on the (d, r) plane for which the traffic gain T_G is constant and equal to b. The curves $r(d, b = 1)$ represent the values of r for which the traffics carried by the TW and TG techniques are the same. It is worth noting that, T_G is an increasing function of r; in fact, the expressions (6) and (15) show that an increase of r entails an increase of $\tilde{C}L_{worst}^{TW}$, whereas the value of $\tilde{C}L_{worst}^{TG}$ remains constant since it is independent of r. As a consequence, in the (d,r) plane, the region above a $r(d, b)$ curve is the region in which $T_G > b$, i.e. where the TG technique presents a throughput gain with respect to the TW technique greater than b. On the contrary, the region below a $r(d, b)$ curve corresponds to a region in which $T_G < b$, so in this region the TG techniques presents a gain with respect to the TW technique lower than b.

The iso-gain curves $r(b, d)$ can be determined by equalizing the traffic gain T_G to b and by solving this equation for r. Since the actual expression of the traffic gain is not available, the upper and lower bounds given by (23) have to be used. Remembering that T_G is an increasing functions of r, it is easy to understand that the solution of the equation $T_G(up) = b$ provides a lower bound of $r(d, b)$, i.e. $r_{low}(d, b)$; on the contrary, for opposite reasons, the solution of the equation $T_G(low) = b$ provides an upper bound of r, i.e. $r_{up}(d, b)$.

In formula,

$$r_{low}(d, b) \le r(d, b) \le r_{up}(d, b) \quad (24)$$

If all of the network links have the same length, the normalized bottleneck link loads bounds can be written as follows:

$$\tilde{C}L_{worst}^{TW}(up) = \sum_{i \in P_k} a_i^s \cdot (1 + d + 2 \cdot r \cdot h_i) \quad (25)$$

$$\tilde{C}L_{worst}^{TW}(low) =$$

$$= \sum_{i \in P_k} a_i^s \cdot \left\{ \sum_{s = z_{i,k}+1}^{h_i} \left[(1 - l_{worst})^{s-2} \cdot l_{worst} \cdot 2 \cdot r \cdot (s - z_{i,k}) \right] + \right. \quad (26)$$

$$\left. + \left[(1 - l_{worst})^{h_i - 1} \cdot (1 + d + 2 \cdot r \cdot h_i) \right] \right\}$$

$$\tilde{C}L_{worst}^{TG}(up) = \sum_{i \in P_h} a_i^s \cdot (1 + d) \quad (27)$$

$$\tilde{C}L_{worst}^{TG}(low) = \sum_{i \in P_h} a_i^s \cdot (1 + d) \cdot (1 - l_{worst})^{(z_{i,h} - 1)} \quad (28)$$

Assuming $(1 - l_{worst})^x \simeq 1 - x \cdot l_{worst}$, i.e. $l_{worst} << 1$, substituting (25) and (28) in the second of (23), equalizing to b and solving for r, we obtain the following lower bound for r:

$$r_{low}(d, b) = \frac{1}{2 \cdot \sum_{i \in P_k} a_i^s \cdot h_i} \cdot \left\{ (1 + d) \left[b \cdot \sum_{i \in P_h} a_i^s - \sum_{i \in P_k} a_i^s \right] - \right.$$
$$\left. - \left[b \cdot (1 + d) \cdot l_{worst} \cdot \sum_{i \in P_h} a_i^s \cdot (z_{i,h} - 1) \right] \right\} \quad (29)$$

Analogously, substituting (26) and (27) in the first of (23) with obtain the following upper bound for r:

$$r_{up}(d, b) =$$
$$\frac{(1 + d) \left[b \cdot \sum_{i \in P_h} a_i^s - \sum_{i \in P_k} a_i^s \right] + \left[\sum_{i \in P_k} a_i^s \cdot (h_i - 1) \cdot l_{worst} \cdot (1 + d) \right]}{\sum_{i \in P_k} a_i^s \cdot \left\{ \sum_{s = z_{i,k}+1}^{h_i} \left[2 \cdot (1 - l_{worst}) \cdot l_{worst} \cdot (s - 2) \cdot (s - z_{i,k}) \right] + \left[2 \cdot h_i \cdot (1 - (h_i - 1) \cdot l_{worst}) \right] \right\}} \quad (30)$$

for $l_{worst} \ll 1$ the previous expression can be approximated as follows:

$$r_{up}(d,b) \simeq \frac{1}{2 \cdot \sum_{i \in P_k} a_i^s \cdot h_i} \left\{ (1+d) \left[b \cdot \sum_{i \in P_h} a_i^s - \sum_{i \in P_k} a_i^s \right] + \left[\sum_{i \in P_k} a_i^s \cdot (h_i - 1) \cdot l_{worst} \cdot (1+d) \right] \right\} \quad (31)$$

The (29) and (31) provide two usable expressions for the lower and upper bound of *r* if the bottleneck links #*k* and #*h*, in case of the TW and TG techniques, respectively, are known. Unfortunately, while in case of the TG protocol the bottleneck link is invariant with respect to *r*, in case of the TW technique the bottleneck link is dependent of *r*, making hard the evaluation of $r_{up}(d,b)$ and $r_{low}(d,b)$. As a matter of example, let us suppose to set a certain value of *r* for which the TW bottleneck link is #*k*, and, by means of (29) and (31), to compute the values of $r_{up}(d,b)$ and $r_{low}(d,b)$; unfortunately, for these computed values of *r*, the TW bottleneck link could be different of #*k* and therefore the computed values of $r_{up}(d,b)$ and $r_{low}(d,b)$ could not be correct. In order to overcome this problem, an iterative procedure has been defined allowing the TW bottleneck link to be determined. The iterative procedure can be summarized in the following steps, it is to be noted that it has to be applied for the evaluation of the TW bottleneck link for the upper and lower bound of *r*, separately; for sake of brevity the parameters to be utilized in case of the upper bound are indicated in squared brackets.

1. set r=0;
2. compute the $CL^{TG}(low)$ [or $CL^{TG}(up)$] vector;
3. set h equal to the index of the maximum element of $CL^{TG}(low)$ [or $CL^{TG}(up)$];
4. compute the $CL^{TW}(up)$ [or $CL^{TW}(low)$] vector;
5. set k equal to the index of the maximum element of $CL^{TW}(up)$ [or $CL^{TW}(low)$];
6. compute the (29) [or (31)] and set $r = r_{low}(d,b)$ [or $r = r_{up}(d,b)$];
7. re-compute the $CL^{TW}(up)$ [or $CL^{TW}(low)$] vector;
8. set k_1 equal to the index of the maximum element of $CL^{TW}(up)$ [or $CL^{TW}(low)$];
9. if k_1 is equal to k then stop the recursion, else set $k= k_1$ and repeat the procedure from the step 6.

When the recursion ends, the values #k and #h represents the TW and TG bottlenecks links to be used in the (29) [or (31)] for the evaluation of $r_{low}(d,b)$ and $r_{up}(d,b)$, respectively.

IV.A. Symmetric Networks

To easily obtain some interesting conclusions, in the following we discuss the numerical results that are obtained assuming that the bottleneck links in case of TW and TG techniques coincide (i.e. *k*=*h*). This assumption strictly holds only for some particular networks, that we call *symmetric networks.* However, in our work, several non *symmetric networks* has been also analyzed and analogous conclusions have been reached. The details of these studies are not reported here for lack of space.

In a *symmetric* network, the (29) and (31) bound the iso-gain curve $r(d, b=1)$ as follows:

$$0 \le r\ (d, b=1) \le \frac{(1+d) \cdot l_{worst}}{2} \cdot \left[1 - \frac{\sum_{i \in P_k} a_i^s}{\sum_{i \in P_k} a_i^s \cdot h_i} \right] \quad (32)$$

while, for $b \gg 1$, as the second squared brackets of the numerators of (29) and (31) can be neglected, the iso-gain curves have the following expression:

$$r_{low}(d, b \gg 1) = r_{up}(d, b \gg 1) = r\ (d, b \gg 1) = \frac{(1+d) \cdot (b-1)}{2} \cdot \left[\frac{\sum_{i \in P_k} a_i^s}{\sum_{i \in P_k} a_i^s \cdot h_i} \right] \quad (33)$$

The (32) and (33) show that:

i) for values of T_G near to 1, *r* is proportional to the value of the loss probability on the bottleneck link l_{worst};
ii) for values of T_G much greater than 1, *r* is instead proportional to (b-1)/2.

As a remarkable conclusion, a general criterion for the optimum choice of the burst size can be given. In fact, by observing that: i) fixed the absolute value of the link length, the shorter the bursts are (i.e. the higher the value of *r* is), the higher the traffic gain is; ii) $\delta = d^{\sec} / T_{burst}$ can also be interpreted as the overhead in a burst transmission due to the configuration time of the optical devices; it follows that, on the one hand, T_{burst} must be great enough such that δ be lower than a given threshold α; on the other hand, T_{burst} has to be chosen as low as possible in order to maximize the T_G. Therefore, the optimum choice of T_{burst} is $T_{burst}=d^{sec}/a$.

IV.B. Ring network

In this section a network with a ring topology is investigated, the assumptions here considered are: i) five node unidirectional ring; ii) minimum hops routing; iii) all of the links have the same absolute length r^{km} in km; iv) only point-to-point traffic; v) uniform traffic, i.e. the offered traffic between every pair of nodes is the same; vi) the bit rate of a wavelength is equal to 2.5 Gbps.

Under the previous assumptions a ring is a symmetric network.

The Fig. 4 shows the behavior of the iso-gain curves for some values of T_G and l_{worst}. It is easily to notice that for T_G=1 the curves are proportional to l_{worst}, whereas for slightly greater values of gain (i.e. T_G>1.2), they grows up towards (T_g-1)/2.

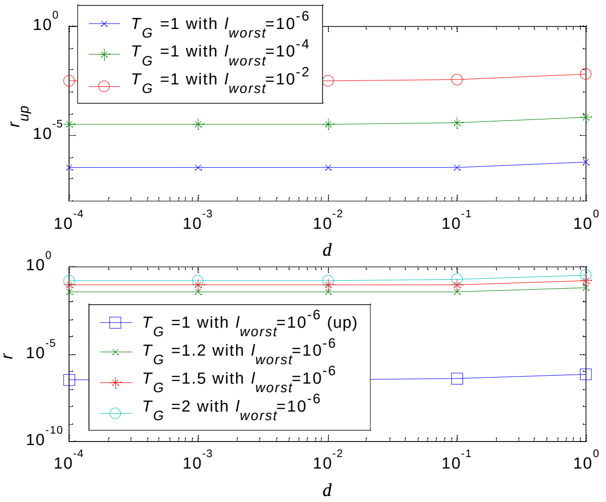

Fig. 4: Unidirectional five node ring with all-to-all traffic distribution iso-gain curves on (d , r) plane.

To give an idea of the *traffic gain* dependence on the network size, we compare TG and TW in two different cases: a) "slow optical devices", that is $d^{\sec}$ = 1 ms ; b) "fast optical devices", that is $d^{\sec}$ = 1μs.

In order to have a value of overhead, given by δ, less than or equal to 0.03, we set the mean burst size (*MBS*) equal to 10 Mbytes in the first case and to 10 Kbytes in the second one.

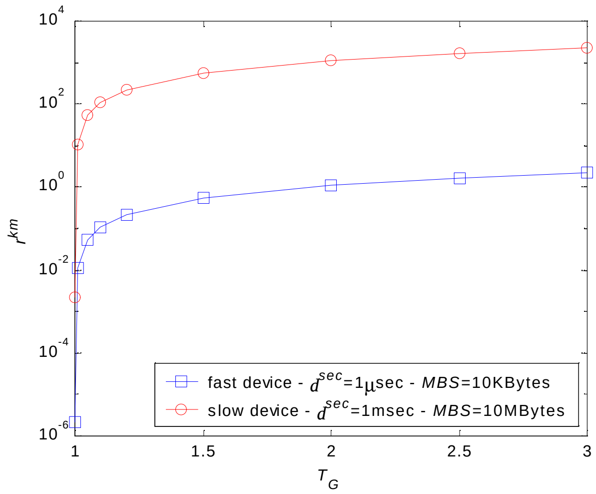

Fig. 5: Unidirectional five node ring, link length (km) vs. Traffic gain, for bottleneck blocking probability $l_{worst} = 10^{-6}$

The Fig. 5 displays the value of not-normalized link length r^{km} versus the *traffic gain* for $l_{worst} = 10^{-6}$; it is worth to noting that, on the basis of the final observation of section IV.C, this curve substantially holds for all the values $l_{worst} \le 10^{-3}$. We notice that small values of gain are already obtained for very short links, whereas, for greater values of gain, the trend rapidly reaches the value given by (33).

So, with slow optical devices, reasonable gains (T_G>1.5) can be achieved only for hundreds of kilometers long links (i.e. wide area networks), whereas, with fast optical devices, the same gain values hold even in local or metropolitan area networks, i.e. for links of length in the range of a kilometer.

IV.C. vBNS

In this section we apply the proposed methodology to a real network represented by the vBNS (very high bandwidth network service). The topology and the traffic matrix of this network have been derived from [14]; the length of the links has been fixed equal to the geographical distance between the nodes. The bit rate over each wavelength has been fixed at 2.5Gb/s and a shortest path routing algorithm has been used. As in the ring network case, we fixed the $l_{worst} = 10^{-6}$ and the TG and TW solutions have been compared in two different cases: a) "slow optical devices", that is $d^{\sec}$ = 1 ms ; b) "fast optical devices", that is $d^{\sec}$ = 1μs.

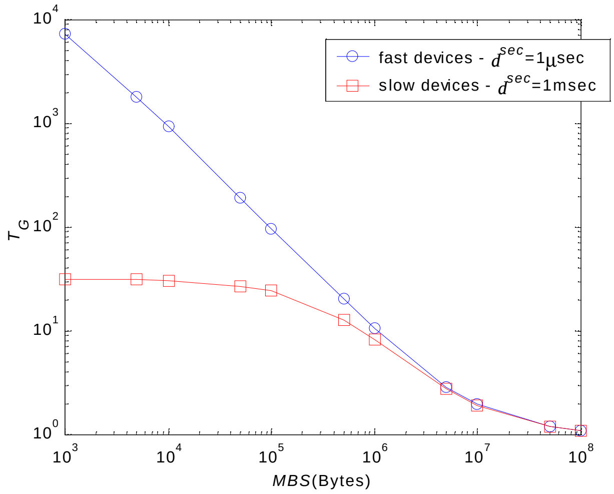

Fig. 6: Traffic gain vs. the mean burst size (MBS) in the vBNS backbone

The Fig. 6 displays T_G as function of the mean burst size. As we expected, due to the long distance covered by the links, there is a clear advantage in the use of the TG solution in both a) and b) cases. In fact, the gain begins to be small only for very large bursts that require large time to be formed and so an unacceptable delay for the IP traffic.

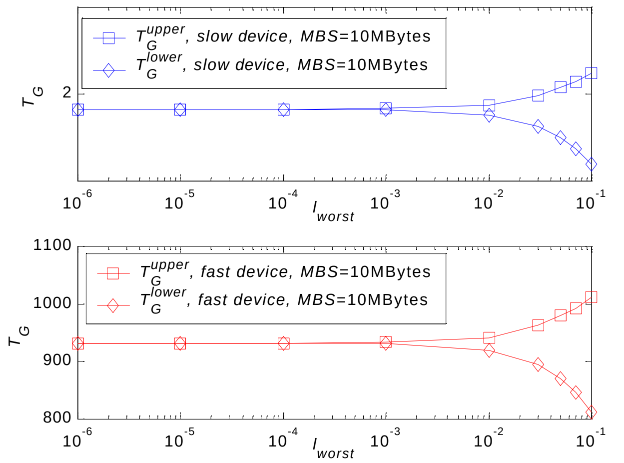

Fig. 7: Traffic gain upper and lower bound vs. l_{worst} in the vBNS backbone

The Fig. 7 shows that for bottleneck link loss probabilities below 10^{-3} the upper and the lower bounds of T_G are practically equal and, due to the high gain environment, they loss the dependence on l_{worst}.

V. CONCLUSION

In this paper a comparative methodology aims at comparing the *Tell and Wait* and *Tell and Go* resource reservation strategies in an optical burst switched network have been proposed. In particular, the comparison of the two techniques has been carried out to determine the relevant convenience regions taking into account: i) the constraints imposed by the optical technology; ii) the network dimensions and iii) the burst sizes. The figure of merit assumed for the comparison is the *traffic gain*, that is the ratio between the amounts of external traffic offered to the backbone, in the TW and TG cases, that gets the same value of worst link blocking probability. The main conclusions of the work are that in networks planned so as to obtain small burst loss probability, with slow optical devices (configuration time more than 1 ms), the Tell and Go significantly increases the network throughput (>50%) only in wide area networks (link length of hundreds of km); whereas, for fast optical devices (configuration time more than 1 μs), this gain can be obtained also in local or metropolitan area networks (link length of the order of km).

VI. REFERENCES

[1] R. Ramaswami, K.N. Sivarjan, "Optical Networks", Morgan Kaufmann Publishers, 1998

[2] M.O'Mahony, D. Simeonidou, S.R. Barnes, "The Integration of Optical Circuit and Packet Switched Transport Networks: a perspective on future evolution", in Ilotron White Papers, http://www.ilotron.com/product.htm

[3] J. Turner, "Terabit Burst Switching", Journal of High Speed Networks (JHSN), Vol.8, No.1, 1999, pp. 3-16

[4] C. Qiao, M. Yoo, "Choices, Features and Issues in Optical Burst Switching ", Optical Networking Magazine, Vol..1, No.2, April 2000, pp. 36-44

[5] C. Qiao, M. Yoo "A Novel Switching Paradigm for Buffer-less WDM Networks", Proceedings of Optical Fiber Communication Conference (OFC), Paper ThM6, Feb. 1999, pp.177-179

[6] S. Verma, H. Chaskar, R. Ravikanth, "Optical burst Switching: A Viable Solution for Terabit IP Backbone", IEEE Network, Vol. 14, No. 6, Nov. - Dec. 2000 , pp. 48-53

[7] Y. Xiong, M. Vandenhoute, H. C. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks", IEEE Journal on Selected Area in Comm. (JSAC), Vol.18, No. 10, Oct. 2000, pp. 1838-1851

[8] G. C. Hudeck, D.J. Muder, "Signaling Analysis for a Multi-Switch All-Optical Network", in Proceeding on Int'l Conf. On Communication (ICC), June 1995, pp. 1206-1210

[9] I. Widjaja, "Performance analysis of burst admission control protocols", IEE proceeding – communications, vol. 142, Feb. 1995, pp. 7-14

[10] V. Paxon, S. Floyd, "Wide area traffic: the failure of poisson modeling", IEEE/ACM Transaction on Networking, vol. 3, no. 3, 1985, pp. 226-244

[11] A. Ge, F. Callegati, L.Tamil, "On Optical Burst Switching and Self-similar Traffic", IEEE Commun. Letts, Vol. 4, No. 3, Mar. 2000, pp. 98-100

[12] T. Doshi et al., "A Simple Data Link Protocol for Next Generation Packet Network", ", IEEE Journal on Selected Area in Comm. (JSAC), Vol.18, No. 10, Oct. 2000, pp. 1810-1823

[13] "FAST ELECTRO-OPTIC DEVICES FOR NEXT GENERATION CROSS-CONNECTS", http://thor.phys.psu.edu/ngi/technologies.htm

[14] J. Bannister, J. Touch, A. Willner – "How Many Wavelengths Do We Reaaly Need? A Study of the Performance Limits of Packet Over Wavelengths", Optical Networking Magazine, Vol..1, No.2, April 2000, pp. 17-28

[15] " Architecture of optical transport networks", ITU-T Recommendation G.872