Adaptive Routing Algorithms for Optical Packet Switching Networks

M. Klinkowski, F. Herrero, D. Careglio and J. Solé-Pareta
Universitat Politècnica de Catalunya
Advanced Broadband Communication Centre – CCABA
Jordi Girona 1-3, 08034 Barcelona
E-mail: {mklinkow, fherrero, careglio, pareta}@ac.upc.es

Abstract—**In this paper, we address the problem of adaptive routing in connection-oriented optical packet switched core networks. We propose several strategies that due to dynamic path selection balance the traffic load inside the network (i.e., implementing traffic engineering functionalities). In addition, proposed algorithms operating in per-packet mode introduce contention resolution capabilities in space domain. Both these functionalities reduce packet loss probabilities and improve network performance characteristics outperforming the conventional static approach based on the shortest path. We present the principles of the algorithms as well as performance results evaluated for recently proposed optical network topologies, namely the NSFNET, the USA backbone network and the EON, the European Optical Network. Finally, using obtained results we detect network bottleneck and redesign the NSFNET network what considerably improves its overall network performance.**

Index Terms—**Adaptive routing, congested link, contention resolution, optical packet switching.**

I. Introduction

Developing optical networks with Quality of Service and Traffic Engineering capabilities is the main goal for the next generation telecommunication infrastructures. An intense debate has been ongoing about which is the optical network model to adopt, aiming at identifying the degree of optical transparency to be achieved, and the proper flexibility of optical interconnection [1]. In the perspective of network optimization, the implementation of packet switching techniques directly in the transport network will bring more statistical sharing of the physical resources to reduce the connection costs. In this direction, two different approaches have been currently developing in the research community: Optical Burst Switching (OBS) and Optical Packet Switching (OPS) [2].

In this paper we focus on the OPS. OPS is presumed to be a long-term solution which requires very fast all-optical switches (switching time in the order of nanoseconds), and advanced optical components such as tunable wavelength converters and 3R optical regenerators. Since OPS is based on statistical multiplexing, packet contentions may arise at the nodes. Therefore, a contention resolution policy must be applied to reduce the packets losses and make the statistical multiplexing more efficient. In the electrical packet switching, contention resolution techniques typically exploit the space domain, by means of deflection routing, and the time domain, by means of queuing. In the optical packet switching, the lack of optical RAMs imposes the use of a pool of fiber delay lines (FDLs) which are bulky and not scalable and offer limited buffering capabilities (few tens of delays at maximum [3]). In contrast, the use of WDM links and wavelength converters allows to solving contention also in frequency domain, by means of wavelength multiplexing.

The design of the contention resolution policy strongly depends on the packet format and the network operation. OPS has usually been studied with reference to fixed length packets and synchronous operation. It permits easier switching matrix design [4] but requires optical synchronization and has limited interworking ability with data-oriented traffic, such as IP. For these reasons recently solutions adopting asynchronous operation and variable length packets have been investigated.

In this paper we focus on the latter. Past works have extendedly studied the contention resolution problem applied to a single switch [5] [6]. In this case, wavelength and time resources can be used as combined tools to reduce the packet loss and, at the same time, to exploit the optical technology features. In [5], the author observes that the FDL buffers are only able to provide discrete delays and this creates gaps between queued packets that can be considered equivalent to an increase of the packet service time, meaning an artificial increase in the traffic load. It has been demonstrated in [6] that a void filling algorithm (called VOID) that aims at minimizing those gaps gives best performance with respect to other policies. Nonetheless, the computational complexity of the VOID algorithm is very high since it requires to knowing the length and the duration of every gap in the queues. A simplification of this algorithm called MINGAP is proposed in [7]; it selects the FDL with the minimum gap only between the last queued packet and the new one.

The next step in the designing of such networks is the study of a global optical network scenario. On this stage, there is a need to introduce routing functions to properly route the optical packets through the nodes and links. In this paper, we address this problem. Our scope is to design routing procedures able to outperforming the conventional static approach based on the shortest path. In particular, we analyze different ways of treating the packets in the network investigating different adaptive (dynamic) strategies for the path selection problem as a function of the local node congestions and global network topology. Concerning adaptive routing in OPS core networks some work has been done in [8], where the authors investigate the effectiveness of statistical multiplexing of packets on different sets of wavelength paths. In contrast, we focus on the examination of various adaptive path selection strategies in order to find the most appropriate for OPS core networks.

The rest of the paper is structured as follows. In section II a discussion on the considered scenario is presented focusing on both the node and network architectures. In section III three different adaptive routing algorithms for path selection are proposed. In section IV, we compare their performance in terms of packet loss probability by using a simulation tool. Finally, Section V concludes the paper.

II. Network Scenario

The considered OPS network comprises several nodes connected in a mesh topology which use an MPLS-like traffic management facility. Packets coming from client layers are classified into a finite number of subsets, called *Forwarding Equivalent Classes* (FECs), based on identification address and quality of service requirements. Each FEC is identified by an additional *label* added to the packets. Edge nodes setup unidirectional connections throughout the network called *Label Switched Paths* (LSPs). Packets belonging to the same FEC are identical from a forwarding point of view and are transferred from source to destination along the LSP which corresponds to their label. On each core node, a simple label matching operation is performed on a pre-computed LSP forwarding table, thus simplifying and speeding up the forwarding function.

Each node (either edge or core) is an optical packet switch with full connectivity and wavelength conversion as shown in Fig. 1. We assume that it is capable to switch asynchronous, variable-length packets [9] allowing a better interworking with heterogeneous client traffic [6]. The switch acts as an output queuing switch using a feed-

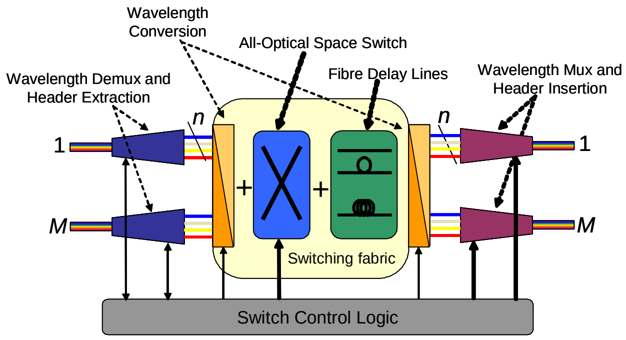

Fig. 1. Node architecture.

forward configuration [10] with the optical buffer made of *B* FDLs. The electronic Switch Control Logic (SCL) takes all the decisions regarding the configuration of the hardware to realize the proper switching actions. When a packet arrives and the label is extracted, the SCL performs the following functions:

1. Lookups the forwarding table to determine the output port *p*, determining also the network path;
2. Searches for the set of wavelength $\Lambda \in p$ not busy;
    a. If $\Lambda = ø$, then the packet is lost;
    b. If Λ ? ø, selects the wavelength $\lambda \in \Lambda$ which introduces the minimum gap with respect to the last queued packets on λ;
3. Determines the delay D_j and selects the FDL *j*;
4. Transmits the packet to FDL *j* with wavelength λ on port *p*.

While the switches implement the MINGAP algorithm to solve the packet contention in time and wavelength domains, the routing algorithm can be exploited to both solve the contention in space domain and balance the traffic load inside the network (i.e., implementing Traffic Engineering (TE) functionalities).

III. Routing in OPS Networks

The next step after definition the node level operations is designing the traffic management rules on the network level. In order to send the data from node A to node B, we need to establish a LSP path between them. It is obvious that between nodes A and B we can trace different routes, so we have to select one of them. The main problem is that a global state of the network is not known when we try to decide which way is a better one. Moreover this factor arises when we assume immense amount of information transported every moment through the OPS network (inaccuracy global network state problem [11]). Therefore we have to define some designing strategies, which attempt to minimize this problem. The first one consist of an application of some global view of the network and

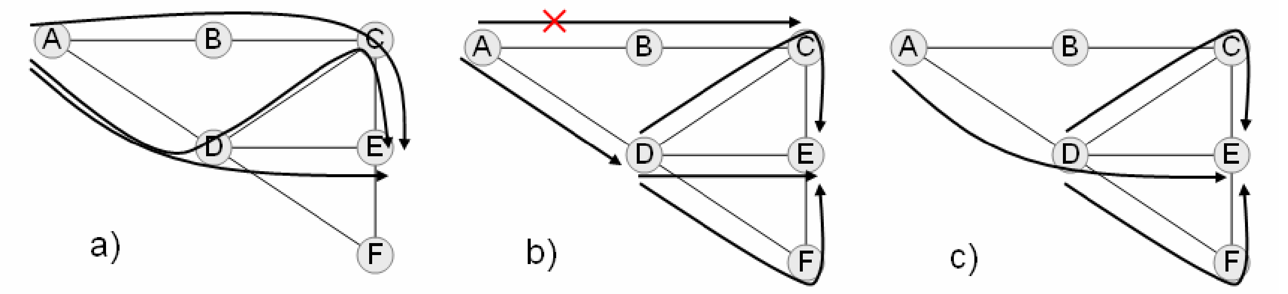

Fig. 2. Example of the packet route using the proposed routing algorithms with k=3,
a) Path excluding (PE), b) Multiple choice (MC), c) By-pass (BP).

establishment multiple LSPs for each pair of the nodes. These multi-LSPs are established only on knowledge of the network topology and a selection of the *k*-shortest paths. Since it is possible to select one path from the set of *k* available, each node can make a decision according to the state of its links. Similarly, while packets traveling through the network, the intermediate nodes can vary the original routes according to some local parameters such as queue congestion or link failure. This strategy provides a better flexibility but in principle keeps the packets more time in the network increasing global network load. Presented strategy has been applied for designing and implementation following three routing algorithms.

The first algorithm is called Path Excluding (PE). The behaviour of this algorithm is the following: each node always selects the less congested output queue among all the ports included in the set of available paths. This selection determines the next hop and excludes from the set of available paths all those paths that not include this hop in their route. Hence, from the *k* original paths, each node removes some paths as long as remains only one path. Figure 2a shows an example. A packet is generated in node A and wants to go to node E. *k*=3 paths are setup: the shortest one is A-D-E, then A-B-C-E and A-D-C-E. If less congested queue of A is on the output port to B, A selects the A-B-C-E path definitely excluding the other possibilities. This means that the rest of the nodes in the selected path cannot take other routing decisions. If the less congested queue of A is on the port toward D, both A-D-E and A-D-C-E are selected while the other is removed. The next node D will take the path decision in the same way. If the less congested queue is toward node E, it chooses the path D-E; otherwise D-C-E is selected. The last case is when all queues have the same congestion level; in this case the node selects the shortest path. It is evident that when all queues are congested, the packet is lost.

The second algorithm is called Multiple Choice (MC). In this case, for each packet each node selects the best path between *k* available paths, with a condition to avoid loops.

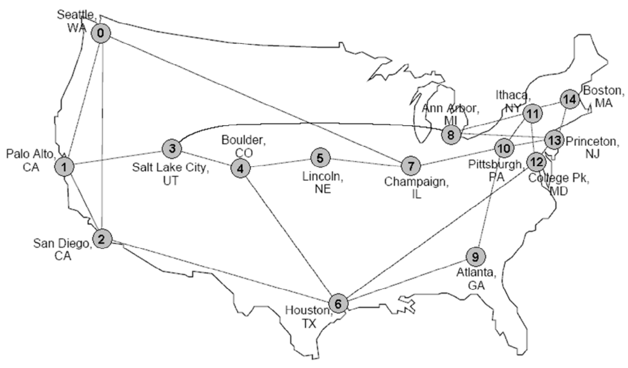

Fig. 3. The NSFNET network topology.

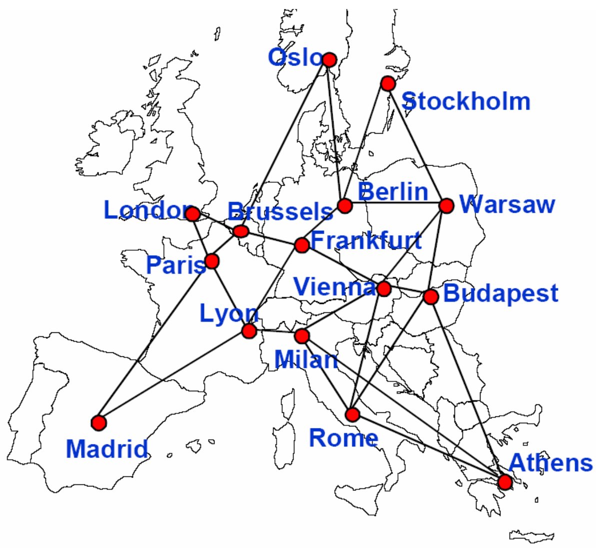

Fig. 4. The EON network topology.

This solution offers a little better flexibility due to the fact that consecutive nodes are not limited in avoiding congested links. But at the same time, the packets can stay in the network for long time, risking an undesirable increase of the global load. Figure 2b shows an example of the algorithm behaviour. As in the previous case, node A wants to transmit a packet to node E using one of the three possible paths: A-D-E, A-B-C-E, and A-D-C-E. In this example, node A selects node D because this output port presents the less congested queue. When packet arrives to D, it can again select one of its three available paths: D-E, D-C-E, and D-F-E. And so on. To limit the time the packets stay in the network, a maximum number of hops (MAXHOP) is defined for this algorithm.

The third algorithm is called By-Pass (BP). For each packet, the source node selects a single path as a function of the state of its output queues. The route can be modified only when traveling packet finds a congested link. In this case, the node tries to 'by-pass' it using an intermediate node to reach the next hop. Figure 2c shows an example of this algorithm behaviour. Node A transmits a packet to node D with destination node E (the path is A-D-E). When packet arrives to node D, no resources are available to reach node E. Therefore, node D finds two by-pass paths in its forwarding table: D-C-E, and D-F-E. It selects the one with the less congested queue.

To compare presented strategies, we also consider a quasi-ideal (QI) algorithm as a benchmarking reference. In this case, we assume that each source node knows in every moment all the information regarding the network congestion situation (the inaccurate problem is not present). Thanks to this knowledge, optimal end-to-end path can be selected for each packet as a function of the state of all links that particular packet can travels through. This path is chosen by the source node and cannot be modified at the intermediate nodes.

IV. Numerical Results

A. Simulation Scenario

In the following section, we present the configuration parameters of the simulated network.

We consider two network topologies:

- the NSFNET (Fig. 3), an USA backbone network, with 15 nodes and 23 bi-directional WDM links
- the EON network (Fig. 4), an European Optical Network, with 15 nodes and 26 bi-directional WDM links.

Both topologies have 15 nodes, but their distribution is quite different. While the nodes in the NSFNET are uniform distributed along the network extension, the EON has a concentration of nodes and links in the middle. Utilizing these networks, we can obtain the algorithms behavior results in two quite different scenarios. For both topologies, we fix the MAXHOP parameter of the MC algorithm to 7.

Regarding the traffic modeling, each node is an edge node capable to generating packets destined to any other nodes (uniform distribution); the interarrival time of the data packets is exponential distributed with a mean that depends on the network load; the packet size is exponential distributed with average and minimum lengths of 500 and 40 bytes respectively. The number of simulated packets is chosen big enough to reach steady-state results.

For both topologies, we assume that each link multiplexes 16 wavelengths at 10 Gbit/s while each node has B = 4 FDLs and implements the MINGAP as a wavelength and time contention resolution algorithms. The granularity of the FDLs is equal to the average packet lengths, i.e. D = 80 m (500 bytes at 10 Gbit/s).

B. Performance Evaluation

In the following figures, we analyze both the Packet Loss Probability (PLP) and the average number of hops per packet as a function of the overall network load (normalized to the network capacity).

Figure 5 and Fig. 6 plot the overall network PLP for the NSFNET topology and EON topology respectively, comparing the routing algorithms. We can see that, excepting MC, all algorithms present better performance than the simple Shortest Path (SP). In particular the QI algorithm produces the best overall results. Nonetheless, we want to remark that QI is used as a benchmarking reference since it is an unviable solution as explained in Section III. Therefore the BP algorithm shows the lower PLP. A strange behaviour can be observed under MC algorithm; in this case both networks start to drop packets only when the traffic load reaches values above 0.725. It can be due to the fact that this algorithm well distributes the packets over the network as long as no link congestion occurs, and it does not generate packets losses. When a link starts to be congested, the nodes try to find alternative paths distributing the packets to the neighbor nodes. But also these nodes try to find alternative paths to reach the destination and so on until the MAXHOP is reached and the packets are discarded. This causes that the packets remain in the network for a long time, increasing the network load and therefore the packet losses.

Another measure we analyze is the average number of hops of packets obtained for analyzed algorithms (Fig. 7 and Fig. 8 for NSFNET and EON topologies respectively). It is evaluated as the overall number of packets processed at the nodes divided by the number of generated packets. This implies that increasing the packet losses, the average number of hops decreases consequently. Both Fig. 7 and Fig. 8 highlight the strange behaviour of the MC algorithm. When some links become congested (around 0.725 load), the nodes send packets to longer paths as the average number of hops indicates. This situation increases the

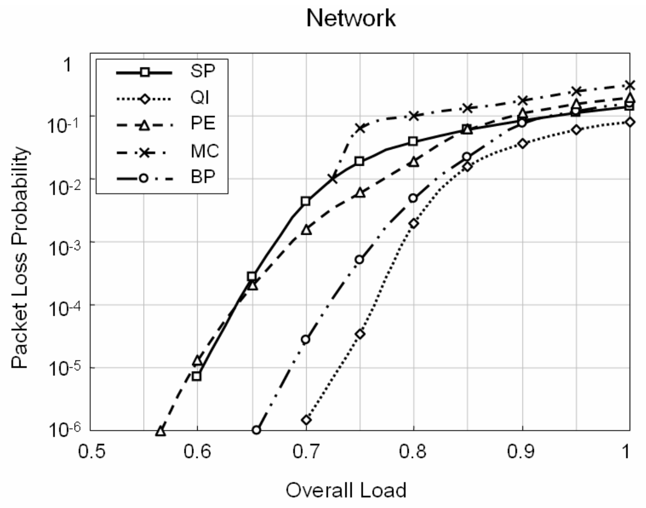

Fig. 5. Packet loss probability as a function of the overall network load comparing the routing algorithms for the NSFNET topology.

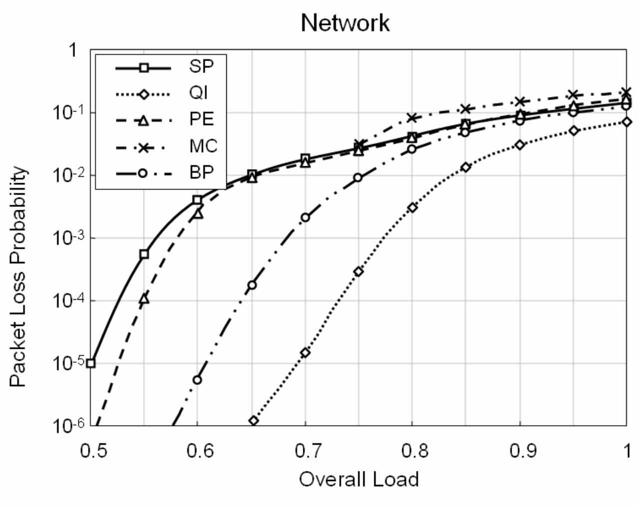

Fig. 6. Packet loss probability as a function of the overall network load comparing the routing algorithms for the EON topology.

network load and therefore the packet losses. Definitely, the MC algorithm is not a good strategy. Again, the simple by-pass strategy presents the better performance reaching the destination in fewer hops.

C. Network Design Considerations

Studying the results obtained from the NSFNET topology, we can draw that the behavior of a global network seems to be dependent on the behavior of a single node. This node is Champaign and in particular the congestion problem resides on the link from Champaign to Pittsburgh, strategically located in the center of the network.

Figure 9 and Fig. 10 show the PLP as a function of the overall load for the Champaign node and the Champaign-Pittsburgh link. We can observe that the tendency is practically the same as on previous graph (Fig. 5) of the overall network. It can indicate the importance of this node for the efficiency of a global network.

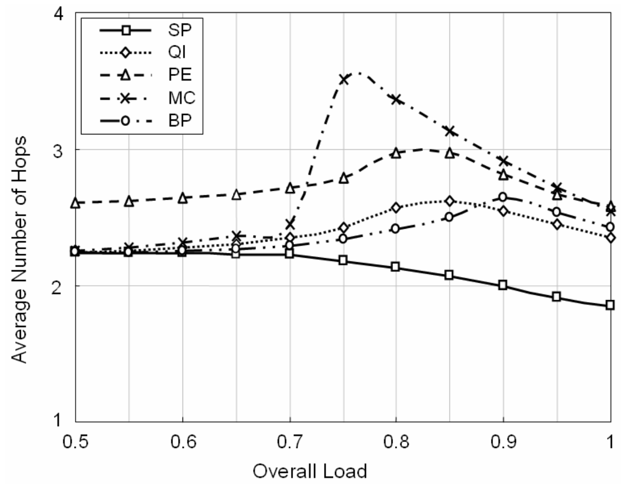

Fig. 7. Average number of hops as a function of the overall network load comparing the routing algorithms for the NSFNET topology.

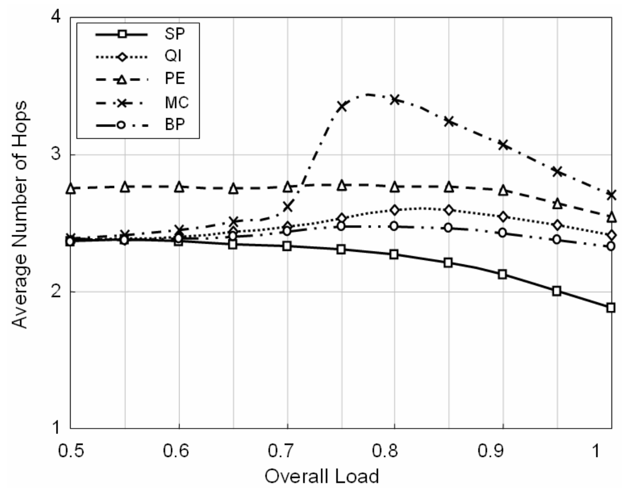

Fig. 8. Average number of hops as a function of the overall network load comparing the routing algorithms for the EON topology.

Taking account this observation, we redesign the network increasing the number of wavelengths on considered link for improving obtained results. Due to the fact that distinct nodes are connected by DWDM links, this task is feasible. Therefore from now on, in place of initial 16 wavelengths, we consider 32 wavelengths on the specific link connecting Champaign and Pittsburgh nodes. The capacity is doubled in both directions.

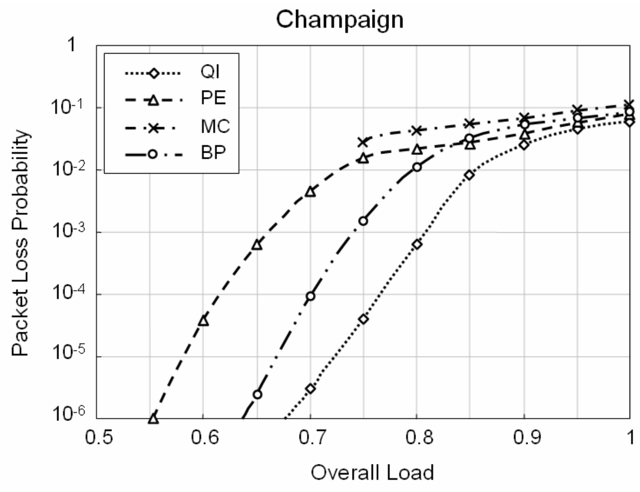

Fig. 9. Packet loss probability as a function of the overall network load comparing the routing algorithms for the Champaign node of the NSFNET topology.

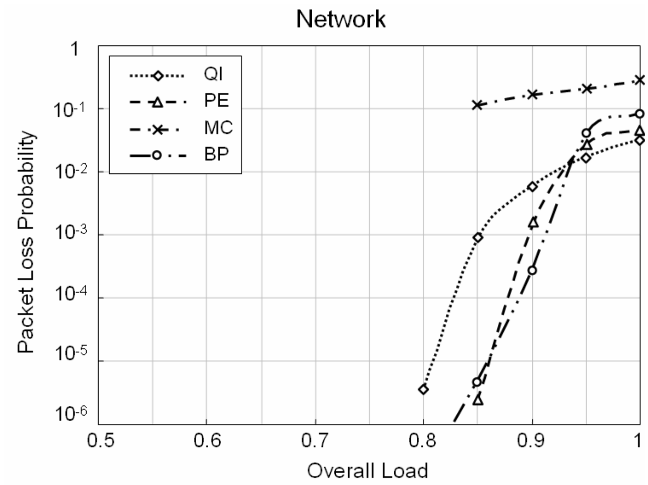

Fig. 11. Packet loss probability as a function of the overall network load comparing the routing algorithms for the NSFNET topology with 32 wavelengths between Champaign and Pittsburgh link.

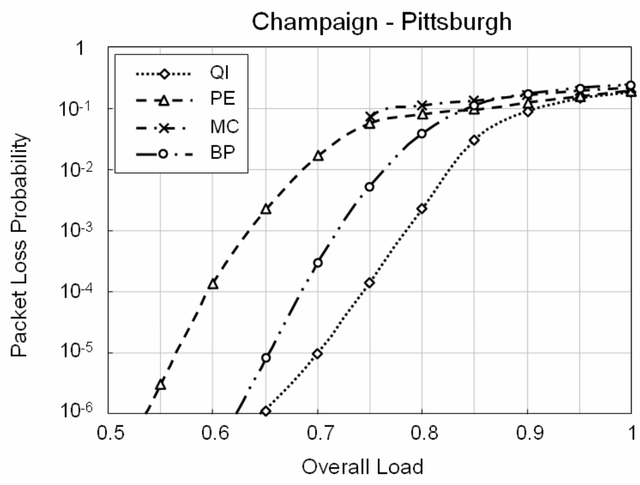

Fig. 10. Packet loss probability as a function of the overall network load comparing the routing algorithms for the link from Champaign to Pittsburgh of the NSFNET topology.

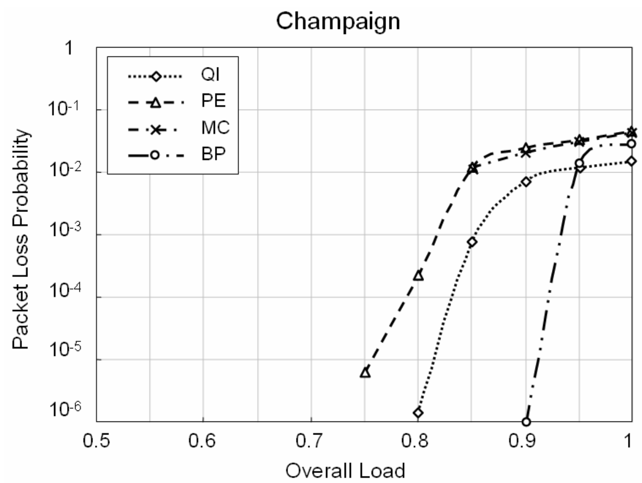

Fig. 12. Packet loss probability as a function of the overall network load comparing the routing algorithms for the Champaign node of the NSFNET topology with 32 wavelengths between Champaign and Pittsburgh link.

In the results (Fig. 11 and Fig. 12), it can be observed that the PLP decreases noticeably, as evidently in the whole network as well as in the node of Champaign. When we compare the Fig. 11 with the Fig. 5 and Fig. 12 with Fig. 9, we observe an evident improvement in the efficiency of the network. In this case, the BP algorithm presents the lower PLP, even better than QI. The MC algorithm is still the worse strategy highlighting again that it is better to limit the number of packet rerouting.

V. Conclusions

In this paper the problem of adaptive routing in optical packet switched core networks has been studied. Different routing strategies, namely the Path Excluding (PE), Multiple Choice (MC) and By-Pass (BP) designed for optical packet networks have been proposed.

The adaptively facility has been introduced by mean of dynamic path selection in a function of the local node congestions and global network topology. Proposed routing algorithms operating in per-packet mode have been exploited to both solve the contention in space domain and balance the traffic load inside the network. Both these functionalities improve network performance characteristics in term of packet loss probabilities.

In particular, simulation results have demonstrated that the reference Quasi-Ideal (QI) algorithm offers the best benefits for both studied network topologies. However, this algorithm is unrealizable due to the fact that it needs accurate global network state information. Hence, for real network scenario, BP algorithm is the best alternative.

Obtained results also helped us in detection the bottlenecks in the network where excessive congestion had been produced. Upgrading overloaded links in additional wavelengths provided remarkable improvement in a global performance results for all considered algorithms.

The main conclusion is that the use of adaptive routing effectively provides TE mechanisms for traffic distribution that lightens overloaded links and nodes in the network. At the same time, the adaptive scheme must be applied only when strictly necessary trying to not increase the time the packets remain in the network. This is the purpose of the BP algorithm and indeed it presents best performance.

Future works will continue the investigation of per-flow adaptive schemes focusing on by-passing not only links but also zones of congestion according to a prediction mechanism.

ACKNOWLEDGMENT

The work has been partially funded by the Spanish Ministry of Science and Technology (MCYT) under contract FEDER-TIC2002-04334-C02-02).

REFERENCES

[1] N. Ghani, S. Dixit, T.-S. Wang, "On IP-WDM integration: a retrospective", *IEEE Commun. Mag.*, vol. 41, no. 9, Sep. 2003, pp. 42-45.

[2] L. Xu, H.G. Perros, G. Rouskas, "Techniques for optical packet switching and optical burst switching", *IEEE Commun. Mag.*, vol. 39, no. 1, Jan. 2001, pp. 136-142.

[3] T.S. El-Bawab, J.-D. Shin, "Optical packet switching in core networks: between vision and reality", *IEEE Commun. Mag.*, vol. 40, no. 9, Sep. 2002, pp. 60-65.

[4] P. Gambini et al., "Transparent optical packet switching: network architecture and demostrators in the KEOPS project", *IEEE J. Select. Areas Commun.*, vol. 16, no. 7, Sep. 1998, pp. 1245-1259.

[5] F. Callegati, "Optical buffers for variable length packets", *IEEE Commun. Let.*, vol. 4, no. 9, Sep. 2000, pp. 292-294.

[6] L. Tancevski, S. Yegnanarayanan, G. Castañon, L. Tamil, F. Masetti, T. McDermott, "Optical routing of asynchronous, variable length packets", *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, Oct. 2000, pp. 2084-2093.

[7] F. Callegati, W. Cerroni, G. Corazza, "Optimization of wavelength allocation in WDM optical buffers", *Opt. Network Mag.*, vol. 2, no. 6, Nov. 2001, pp. 66-72.

[8] F. Callegati, G. Muretto, C. Raffaelli, P. Zaffoni, W. Cerroni, "Adaptive routing in DWDM optical packet-switched networks", in *Proc. 8th Opt. Network Design Model.* (ONDM 2004), Gent, Belgium, Feb. 2004.

[9] D. Chiaroni et al., "First demonstration of an asynchronous optical packet switching matrix prototype for multiterabit-class routers/switches", in *Proc. 27th Eur. Conf. Opt. Commun.* (ECOC 2001), Amsterdam, The Netherlands, Oct. 2001.

[10] D.K. Hunter, M.C. Chia, I. Andonovic, "Buffering in optical packet switches", *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 12, Dec. 1998, pp. 2081-2094.

[11] X. Masip-Bruin, S. Sànchez-López, J. Solé-Pareta, J. Domingo-Pascual, "QoS routing algorithms under inaccurate routing information for bandwidth constrained applications", in *Proc IEEE Int. Conf. Commun.* (ICC 2003), Anchorage, AL, May 2003, pp. 1743-1748.