Integrated Routing and Grooming in GMPLS-Based Optical Networks

Ding Zhemin and Mounir Hamdi
Department of Computer Science
Hong Kong University of Science and Technology

Jack Y. B. Lee
Chinese University of Hong Kong
Victor O. K. Li
University of Hong Kong

Abstract—**This paper proposes an integrated routing and grooming algorithm for IP over WDM networks. Assuming a peer model in GMPLS-Based optical networks, we take into account the combined topology and resource usage information on both IP and WDM layers. Based on a clustering technique called Blocking Island Paradigm, we propose an enhanced Blocking Island Graph (BIG) network model with Blocking Island Hierarchy (BIH) to abstract network resources. The main idea of the algorithm is to keep the integrity and load balance of related Blocking Islands. We also combine a cost function in the routing algorithm to groom traffic flows into active lightpaths. The complexity of the algorithm is analyzed to show its efficiency. In the simulation, we compare the algorithm with three other integrated routing algorithms in terms of blocking probability. The three algorithms are: the integrated min-hop (IMH) routing algorithm, the maximum open capacity (MOCA) routing algorithm and the IP-WDM grooming (IWG) algorithm. Simulation results show our algorithm has the best performance.**

Keywords-GMPLS; Routing; Grooming; IP over WDM

I. INTRODUCTION

The integration of IP over WDM involves many issues. One of them is the need to fast and efficiently deploy bandwidth guaranteed end to end paths between ingress-egress routers. Motivated by this need, we develop an integrated routing and grooming algorithm for IP over optical networks which takes into account the combined knowledge of two layers and aggregate different traffic flows to efficiently use the network resource. This integration scheme is more robust and flexible to changing traffic than previous schemes which only do dynamic routing in the IP layer and use a static wavelength routed network topology underneath. Notice the two layer-routing without good coordination is prone to several problems. One of them is the bandwidth of some wavelengths may be occupied by small amount of traffic and severely underused. This situation could be prevented if proper traffic grooming is carried out on the upper layer. On the other side, the congested routers are the main reason for the reduced performance in the IP layer. The congestion of routers in the IP layer is possible to be relieved by adjusting the virtual topology of WDM networks.

In the proposed algorithm, we assume the bandwidth requirement can be a fraction of lightpath and the interconnection topology of two layers can change

dynamically. We also assume that similar control planes are used in the IP and WDM networks. Specially, the recent advances in *generalized multiprotocol label switching* (GMPLS), proposed by the Internet Engineering Task Force (IETF), have provided the enhanced signaling and routing mechanism in both IP and WDM layer, which makes it possible for the implementation of our integrated routing and grooming algorithm.

The remainder of the paper is organized as follows: we firstly introduce different IP over WDM models and present a brief survey on various integrated routing algorithms. In section III, we define the routing problem and network model Based on GMPLS. A paradigm called Blocking Island and an integrated routing and grooming algorithm based on this paradigm are discussed in section IV and V. In section VI, we compare and analyze the simulation results of different routing algorithms. Finally the conclusion of this paper is given.

II. RELATED WORK

A general IP over WDM networking model is shown in Figure 1, where multiple optical subnetworks are interconnected by well defined signaling and routing interface, also known as the network-network interface (NNI), to function as a core optical network. Clients (IP/GMPLS routers) are attached to the optical core network, setting up end to end path over the virtual topology defined by optical lightpaths. The signaling and routing information between clients and core optical networks is interchanged by the user-network interface (UNI).
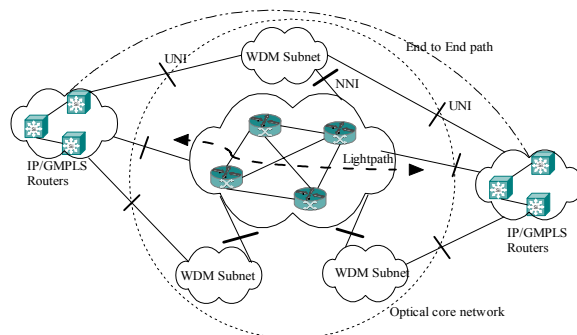

Figure 1. A general IP over WDM networking model

Based on different degree of information sharing and control sharing between two layers, we have three interconnection models [1]: overlay model, augmented model and peer model. In this paper, we study the integrated routing and grooming algorithm under the peer model assumption. We firstly present a brief survey on the available integrated routing algorithms. A framework is proposed in [1] to use Open Shortest Path First (OSPF) with "suitable" optical extensions. The assumption is that link state information is maintained by all nodes and IP routers and OXCs are treated equally in a joint network topology. In this framework, although the optical layer seems transparent to the upper IP layer, the IP routing and lightpath routing are still carried out separately as a two step process, failing to make full use of the combined information. The algorithm proposed in [2] focuses on grooming. By incorporating the information of incoming traffic flows with load distribution of lightpaths, a constraint-based routing algorithm is developed. The main objective of this algorithm is to optimize the resource usage by grooming a new flow into an already existing lightpath. A cost function is designed to measure the possible grooming. It is difficult to find appropriate cost coefficients, which are derived in [2] from experiment results. [3] proposes an algorithm for integrated dynamic routing of bandwidth guaranteed paths in IP over WDM networks. The key idea is to pick paths that don't interfere too much with potential future traffic flows. The interference is estimated by computing the maximum flow between ingress-egress routers. The algorithm takes into account the combined knowledge of resource and topology information in both IP and optical layers. However, the computation of the algorithm is quite expensive and it assumes the information of ingress/egress nodes is already known.

III. PROBLEM DEFINITION AND NETWORK MODEL

In this section, a possible IP over WDM architecture based on GMPLS is presented. Notice in our algorithm, the only dynamic information needed is the bandwidth availability in IP layers and wavelength availability in WDM layers. We define a network topology $G (V, L, W)$ for a given IP/WDM network, where V is the set of all nodes; L is the set of bidirectional optical links and W is the set of wavelength per fiber link. Here we assume the number of wavelength on each fiber link is the same. Under the peer model assumption, network nodes are treated as integrated router/OXC nodes and there is only one control plane. While in practice, it is possible some nodes, which are only OXCs without the function of IP routers, remain one part of an IP/WDM network. Therefore we consider $V(R, O)$ for a given set of nodes in an IP/WDM optical network, where R is the set of integrated router /OXC nodes, O is the set of OXC nodes. Nodes in R can multiplex or demultiplex traffic requests with any bandwidth granularity and do wavelength conversion. Nodes in O can only multiplex or demultiplex traffic requests with the bandwidth granularity of a whole wavelength and don't have wavelength conversion capability. The transaction power of each node is only limited by the network resource availability. This assumption can be relaxed if the detailed transaction capability of equipments is given. An example of network topology is shown in Figure 2, where node 1, 2, 3, 4, 8 are integrated router/OXC nodes and node 5, 6, 7 are plain OXCs.
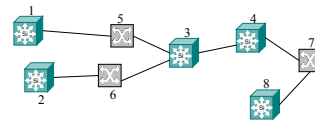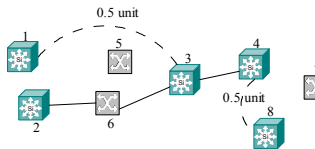

Figure 2.   An example of IP over WDM network

Figure 3.   A new topology of the example network

Based on the GMPLS framework, an optical channel (λ-LSP) needs to be set up for each request and the required bandwidth is reserved on links of the λ-LSP path. The request to set up a λ-LSP can be defined as (X_μ, Y_μ, β_μ) where X_μ and Y_μ are distinct nodes of the network; β_μ is the required bandwidth. Since this is a circuit switched network, the only QoS requirement we consider in this paper is bandwidth. The main objective of our algorithm is to find a bandwidth guaranteed path, either in the existing IP layer or in a new topology by setting up new wavelength paths. Assume the bandwidth of a whole wavelength is 1 unit. A request $(X_\mu, Y_\mu, \beta_\mu)$ is to be routed from node $X_u \in R$ to node $y_u \in R$ with the bandwidth requirement $\beta_u \leq 1$ unit. If an optical channel is set up to reach the destination and this path involves nodes of OXCs, some cut through arcs (lightpaths) may be created to meet the requirement. The IP layer network topology will be changed in this case. For example, in figure 2, a traffic request arrives, requiring the bandwidth of 0.5 unit from node 1 to node 8. To simplify the example, we consider each fiber has only one wavelength. Assume a LSP path (1->5->3->4->7->8) has been found from node 1 to node 8 along the wavelength w_1. Because OXCs can only multiplex and demultiplex traffic requests with the bandwidth request of a whole wavelength, new lightpaths are set up to directly connect integrated nodes. In figure 3, 2 new lightpaths (cut through arcs) are introduced to form a new topology. Notice only 0.5 unit bandwidth is consumed along the path. The residual 0.5 unit bandwidth is still available along the lightpath for future use. Those lightpaths are logical links in the IP layer. They can be released or re-setup according to traffic requests and resource availability.

IV. BLOCKING ISLAND PARADIGM

In this section, we give an introduction on the paradigm used in our routing and grooming algorithm. The Blocking Island paradigm proposed in [4] provides an efficient way of abstracting resource (especially bandwidth) available in a communication network. Blocking Island clusters parts of network according to the bandwidth availability. A β-BI (β-Blocking Island) for a node x is the set of all nodes of the network that can be reached from x using links with at least β available bandwidth. For example, Figure 4 (a) is a network topology of NSFNet where the available bandwidth of each link is given in brackets. $L_n(x)$ refers to a link n which has x

available bandwidth. $N_1 = (V_1, V_2, V_3, V_4)$ is a *40-BI* for node V_1, which means any other node is not able to find a path with at least 40 available bandwidth to reach V_1. Given a traffic request with β (or more) bandwidth requirement, if the request can be satisfied. The source node and the destination node must be in the same β-BI and links of the route (if it exists) must be in the β-BI. However, it is not true that all links inside the β-BI have at least β available bandwidth.
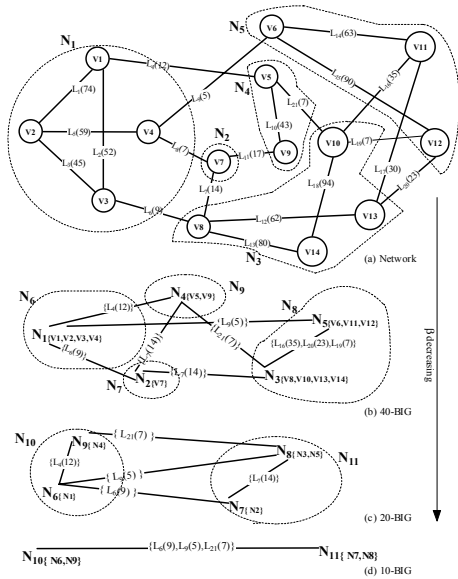

Figure 4. The blocking island hierarchy for bandwidth requirement {*40, 20, 10*}. (a) Network graph; (b) *40-BIG*; (c) *20-BIG*; (d) *10-BIG*

With the abstract technique, instead of studying the whole network topology, we focus our attention only on a small part. For example, given a traffic request (V_1, V_4, *40*) in figure 4, according to the BIH, the route is in *40-BI N_1* if it exists. In the N_1 Blocking Island, different routing heuristic can be employed to find the route. If the bandwidth requirement is not the exact match of the available blocking island graphs, one possible solution is to build a β-BIG (β is the required bandwidth) but it is time-consuming and in some cases, it may not be feasible. So we instead look at the blocking island graph with the closest value. For example, given a traffic request (V_1, V_4, *50*), Based on BI properties, we know the route must be in *40-BI N_1* if it is available. The search space is probably a little bigger than that of a *50-BIG* while it is still much smaller than the whole network topology. If a route is allocated, the available link capacity is decreased and the *BIH* may need to be modified. Based on analysis of the consequences of the splitting, a heuristic called *Minimum Splitting (MS)* is proposed. The difference between this heuristic and the others is that it tries to find a route which doesn't provoke a split in the corresponding BI graph. If the splitting is unavoidable, we would rather select a route which incurs the fewest splitting in the BI graph.

V. INTEGRATED ROUTING AND GROOMING ALGORITHM BASED ON BLOCKING ISLAND PARADIGM

To accommodate those variations, we propose an enhanced BIG network model to represent the IP over WDM network.

We assume a single fiber optical network and the number of wavelengths on each fiber link to be the same. Consider a network topology *G (V, L, W)* for a given IP/WDM network, where *V=(R, O)*, *R* is the set of integrated router/OXC nodes, *O* is the set of OXC nodes; *L* is the set of bidirectional optical links and *W* is the number of wavelengths per fiber link. The enhanced BIG network model can be obtained from a given network topology G as follows. First, we replicate the original topology of *G* |*W*| times. Each copy represents a wavelength and has the same topology as the original IP over WDM network. Then we check each integrated node $r \in R$. In the enhanced BIG model, node r has |*W*| copies as (r_1, r_2 ...$r_{|W|}$). Virtual links r_1r_2, r_2r_3 ... $r_{|W|-1} r_{|W|}$ are then added to connect the corresponding nodes. Notice since those nodes like (r_1, r_2 ...$r_{|W|}$) are actually the same node, the virtual links only characterize the link connectivity with unlimited bandwidth and have no weight (distance) in the routing and grooming computation.

A. Route Existence Check and BI Assigning Procedure

After predefining the proper BIH, when a new traffic request arrives, we pick up the closest BIG level in the BIH to apply routing heuristics. Consider a request $Du= (X_\mu, Y_\mu, \beta_\mu)$ where X_μ and Y_μ are source node and destination node, β_μ is the required bandwidth, using the BI Routing Existence property we immediately know whether the request can be satisfied or not based on a β_μ-BIG without any computing. As we stated before, the given BIH building is not desirable because of time and high maintenance cost. With the predefined limited levels of BIH, it is possible we don't have an exact match of BIG but we can still check the route existence of most requests much faster than a full network search. The route existence screening process is illustrated by an example. Assume a predefined *H* level BIH $(\alpha_1, \alpha_2... \alpha_H)$, where α_i is the bandwidth level of the corresponding BIG level and $\alpha_1 < \alpha_2 < ... < \alpha_H$. If β_μ is equal to any predefined bandwidth value α_i the result can be obtained immediately.

If $\beta_\mu > \alpha_H$, we assign *Du* to α_H-BIG. Then we check whether X_μ and Y_μ are in the same BI of α_H-BIG. If the answer is no, the request is blocked. If yes, we have to do a further check on this BI using Dijkstra's algorithm or a link-state routing protocol.

If $\beta_\mu < \alpha_1$, we assign *Du* to α_1-BIG. Then we check whether X_μ and Y_μ are in the same BI of α_1-BIG. If the answer is yes, the route exists. If no, we have to do a further check on the whole network topology using Dijkstra's algorithm or a link-state routing protocol. This is the worst case in our screening process.

If $\alpha_1 < \beta_\mu < \alpha_H$, say $\alpha_i < \beta_\mu < \alpha_{i+1}$ ($1 \le i \le H-1$), we first check whether X_μ and Y_μ are in the same BI of α_{i+1}-BIG. If the answer is yes, the route exists. If not, we then check whether X_μ and Y_μ are in the same BI of α_i-BIG. If they are in the same BI, we have to do a further check on this BI using Dijkstra's algorithm or a link-state routing protocol. If not, the request is blocked.

Consider all the scenarios, except in the worst case we have to check the whole network topology, normally, we can tell the route existence immediately or only need to do searching in a

much smaller space. By analyzing the traffic statistics and carefully distributing the BI hierarchy, we can reduce the computation cost significantly and identify the bottleneck links more efficiently.

B. Grooming

In the IP/WDM integrated routing, it is better to reuse the active lightpaths than create a new lightpath. Our algorithm dynamically combines different traffic flows to efficiently use the network resources by adding a cost function which contemplates the possibility of reusing an already established lightpath to reach the destination. The cost function is represented by the expression:

Path_Cost=N links + P * Active_Lightpaths

Where N is the number of new optical links along the path; active lightpaths mean the cut through arcs that have already been set up. P is the cost coefficient that can be adjusted to the best network performance. Usually, P is less than one to reduce the total cost of the path and encourage the usage of active lightpaths. The experimental results in [2] show that based on the cost function, dynamic grooming is achieved automatically assuming that the link state database is properly updated. In our simulation, we find even when P is equal to 1, the path composed of active lightpahts is more likely to be picked if shortest path heuristic is employed. The reason is that the cut through arcs which usually bypass several optical links tend to be shorter.

C. Integrated Routing and Grooming Algorithm

In the following, we introduce the steps of the integrated routing and grooming algorithm. The traffic requests arrive dynamically and new lightpaths are set up or released as needed. Firstly the network topology is transformed to the enhanced BIG network model. The bandwidth values of different BIH levels are decided based on previous traffic statistics. The incoming traffic request is assigned to the closest hierarchy level according to the bandwidth requirement. We then compute the path in a blocking island holding both the source node and the destination node. The path satisfying the requirement will be shortest in terms of cost function and causes the least impact on the integrity of the whole network. We then compute K different weighted shortest routes. Each route is assigned to different levels of the BIH to compute the splitting number. The splitting happens in the level with the lower bandwidth value does more damage to the integrity of the network than in the level with the high bandwidth value. Once the path is decided, the enhanced BIG network model is modified with the addition of the lightpaths and BIH is updated with the new bandwidth availability on affected links. Notice the above description is about the request arrival. Once a request leaves, the whole procedure needs to be reversed. The bandwidth of affected links is increased and the BIH is updated as needed. If the available bandwidth of an IP logical link (a lightpath in the optical layer) is 1, the logical link is released and the optical links composed of the corresponding lightpath will be reintroduced to the BIG network model. Below is a formal description of the algorithm. We assume the network has been transformed into an enhanced BIG network with the BIH constructed.

Algorithm:
Input: Dynamic traffic request D with bandwidth requirement β; a BIG network and corresponding BIH
Output: A route satisfying D; if there is no such a route, D is blocked
Description:
1. Assign D to appropriate BIG of the BIH according to β;
2. Route existence check; if the route is not available, D is blocked;
3. Compute K different weighted shortest routes using the path cost function;
4. Assigned each of K routes to all levels of BIH to compute the splitting cost;
5. Order K routes and pick up the route(s) with minimum splitting cost;
6. if there are routes with the same minimum splitting cost, choose one randomly or using other heuristics such as the most loaded link heuristic;
7. Update BIG network and the affected blocking islands in BIH;

D. Complexity analysis

Define a network topology G (V, L, W) for a given IP over WDM network, where V is the set of nodes, L is the set of links and W is the set of wavelengths per fiber link. Assume the set of wavelengths on each fiber link is the same. After transforming into the enhanced BIG network, the number of nodes in the BIG network is $|VW|$ and the predefined BIH level is H. Assume $V=(R, O)$, R is the set of integrated router/OXC nodes, O is the set of OXC nodes. The number of links is equal to $|LW|+|R(|W|-1)|$, $|R(|W|-1)|$ is the number of added virtual links which are regarded as having unlimited bandwidth and no weight. The most common operation in the integrated algorithm is the Blocking Island construction. The β-BI for a given node x of a network can be obtained with a simple greedy algorithm. Starting with an initial set $\{x\}$, we recursively add every node to the set, if this node can be reached from any node in the set by a link that has at least β available bandwidth. In the worst case, this construction process will examine all links. Therefore, the β-BI construction process is linear in $O(n)$, where n is the number of actual links in the network $(n=|LW|)$. The time of constructing one level of BIG is O (mn), where m is the number of nodes and n is the number of links in the network $(m=|VW|, n=|LW|)$. The building time of BIH is O (Hmn). Given a good distribution of BIH levels, usually the route existence check time is equal to the time of constructing one level of BIG. In the worst case, it has to compute the route in the whole enhance BIG network using Dijkstra's algorithm to decide whether the route is available or not. If the request can be satisfied, the running time is equal to the combination of 1) K alternate shortest paths; 2); Splitting cost computation 3) BIH update. That is $K*O(nlg(m))+K*O(Hmn)+O(Hmn),$where K is a constant and H is a constant, so the algorithm running time is linear in $O(mn)$.

VI. PERFORMANCE STUDIES

In this section, we will compare the performance of our routing algorithm with the integrated Min-hop routing algorithm (IMH), the Maximum Open Capacity algorithm (MOCA) [2] and the IP-WDM grooming algorithm (IWG) [3]. The main performance measure we use is the blocking probability, which is the ratio of the number of rejected requests to the number of total arriving requests. We employ a

random dynamic traffic model to generate the incoming traffic. Calls (requests) arrive at each node according to an independent Poisson process with an arrival rate α. An arriving session is equally likely to be delivered to any node in the network. The session holding time is exponentially distributed with mean $1/\mu$. The bandwidth requirement is uniformly distributed between 0.1 and 1 unit. In our simulation, extensive tests are carried out to ensure a steady state is reached. Because our algorithm assumes the bandwidth requirement of requests can be a fraction of a wavelength, we need to modify the incoming traffic model to limit the source and destination nodes within the set of integrated router/OXC nodes. As to other heuristics used in the simulation, the integrated Min-hop routing employs the model proposed in [2] to represent IP over WDM networks and the shortest path heuristic is used to compute the route; MOCA computes the critical weight of all links before applying the weighted shortest path heuristic. Notice the MOCA algorithm must have priori knowledge on the ingress/egress nodes; the IP-WDM grooming algorithm aggregates traffic flows on active lightpaths by using a path cost function similar to the one we use. In the simulation, we assume the cost coefficient P is 0.8.

We tested the performances of the algorithms on NSFNet which has 14 nodes and 21 links. The integrated router/OXC nodes are shaded and the nodes that are not shaded are OXCs with no wavelength conversion. The number of wavelengths per link is assumed to be 3. We predefine the BIH with 0.1-BIG, 0.3-BIG, 0.5-BIG and 0.8-BIG. In figure 5, 5 nodes are shaded. Since any two of those 5 nodes can be ingress/egress node pairs, the number of ingress/egress pairs is 10. Figure 6 shows the simulation results on the network shown in figure 5. Clearly our integrated algorithm performs much better than the other algorithms. MOCA trails behind. IMH and IWG give the worst performance. For example, when the total load is 90 Erlang, the blocking probability for BI is only 3.7×10^{-2}, compared with 7.1×10^{-2} for MOCA, 1.9×10^{-1} for IMH and 2.1×10^{-1} for IWG. The performance of IWG fluctuates around IMH with the system load increasing. The obvious reason is this algorithm focuses mainly on grooming and sometimes may occupy more resources to achieve the grooming.
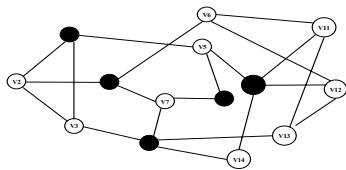

Figure 5. 14 nodes NSFNet with 5 wavelength converters

We also tested our algorithm in other large networks. For example, we tested those algorithms in a network with 100 nodes and 130 links. There are 30 integrated router/OXC nodes. The computation time for MOCA is very long since we have got 870 ingress/egress nodes to calculate. It is not practical for MOCA to handle dynamic traffic in this case. In the rest of 3 algorithms, our algorithm still gives the best performance. Because of the length of the paper, we here don't give the detailed figures.
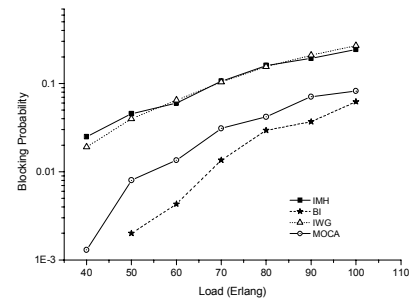

Figure 6. Comparison results of blocking probability for NSFNet

VII. CONCLUSION

The contribution of this paper is to propose a novel integrated routing and grooming algorithm for routing bandwidth guaranteed paths. To the best of our knowledge, it is the first algorithm that considers routing and grooming together in IP over WDM networks. In this paper, we assume a GMPLS-based peer model to implement the algorithm. Borrowing ideas from a clustering technique called Blocking Island Paradigm, we propose an enhanced BIG network model to abstract network resources. We discussed in detail the BIH building, route existence check, BI assignment and grooming cost function. A routing heuristic called Minimum Splitting is then developed. The idea is to keep the integrity of blocking islands as intact as possible. The comparison results in the simulation clearly demonstrate that our algorithm has a very good blocking probability performance. We are currently exploring the feasibility of the distributed implementation of the proposed algorithm to even reduce its time complexity further. We are also working on an extension of this algorithm to address the provision/protection problem in IP over WDM networks.

REFERENCES

[1] B. Rajagopalan *et al*, "IP over Optical Networks: Architectural Aspects," *IEEE Commun. Mag.*, vol. 38 no. 9, Sept. 2000, pp. 94-102.

[2] J. Comellas *et al*, "Integrated IP/WDM Routing in GMPLS-Based Optical Networks," *IEEE Network,* March/April 2003, pp. 22-27.

[3] M. Kodialam and T. V. Lakshman, "Integrated Dynamic IP and Wavelength Routing in IP over WDM Networks," *IEEE INFOCOM2001*, pp.358-366.

[4] C. R. Frei, "Abstraction Techniques for Resource Allocation in Communication Networks," *PhD thesis*, Swiss Federal Institute of Technology – Lausanne, 2000

[5] D. Zhemin and M. Hamdi, "A Simple and Intelligent Routing and Wavelength Assignment Algorithm for All-Optical Networks, "*SPIE Opticom,* 2001, pp. 210-226.

[6] L. Berger, Ed., "Generalized MPLS – Signaling Functional Description, " Internet draft, draft-ietf-mpls-generalized-signaling-09.txt, 2002, work in progress.

[7] J. Moy, "OSPF: Anatomy of an Internet Routing Protocol," *Addison-Wesley*, 1998

[8] A. Banerjee *et al.*, "Generalized Multiprotocol Label Switching: An Overview of Signaling Enhancements and Recovery Techniques," *IEEE Commun. Mag.*, vol. 39, no. 7, July 2001, pp. 144-51