Optical Networks With Hybrid Routing

Hong Huang, *Member, IEEE*, and John A. Copeland, *Fellow, IEEE*

***Abstract*—All-optical switching or wavelength routing has the benefit of optical bypass that can eliminate expensive high-speed electronic processing at intermediate nodes and reduce significantly the cost of high-bandwidth transport. But all-optical switching has the limitations of coarse granularity, lack of multiplexing gain, and scarcity of wavelength resources, which do not mesh well with Internet traffic that has many small and diverse flows and emphasizes the importance of resource sharing. In particular, wavelength routed light paths have difficulty to seamlessly converge with multiprotocol label switching label-switched paths that have arbitrary bandwidth granularity and relatively abundant labels. In this paper, we propose a hybrid wavelength and subwavelength routing scheme that can preserve the benefits of optical bypass for large traffic flows at the same time provide multiplexing gain for small traffic flows. We first study the hybrid routing scheme using static optimization that produces an optimal path set and a partition between wavelength and subwavelength routing. We then present a dynamic heuristic that tracks the static optimization closely. During the process, we proposed a traffic arrival process called incremental arrival with sporadic random termination to more accurately model practical optical network traffic generation process.**

***Index Terms*—All-optical switching, optical networks, subwavelength routing, wavelength division multiplexing (WDM), wavelength routing.**

I. Introduction

ALL-OPTICAL switching is envisioned to be a cost-effective mean to implement future high-speed networks due to its capability to switch traffic end to end in the optical domain and bypass intermediate electrical processing entirely [1]–[4]. Because of such capability, all-optical switching holds the promise to bridge the gap between huge optical bandwidth (on the order of terahertz) and limited electrical processing speed (on the order of gigahertz). In addition, the transparent nature of optical switching makes it inherently multiservice/protocol capable, easing the pain in network protocol changes and transmission speed upgrades.

However, compared with electronic switching, all-optical switching is still immature in functionality. There are a couple of ways to realize all-optical switching, such as optical packet switching, optical burst switching, and wavelength routing. Of the three, wavelength routing is the most mature one with commercial products already on the market. However, wavelength routing has some limitations that do not mesh well with Internet traffic. First of all, wavelength routing is a form of circuit switching, which in itself does not disadvantage it from supporting Internet traffic, especially in view of the current effort within Internet Engineering Task Force (IETF) to introduce multiprotocol label switching (MPLS), also a form of circuit switching, to facilitate traffic engineering and value added services such as VPN. What sets wavelength channels apart from MPLS label switched paths, or ATM circuits for that matter, is that wavelength channels have large and fixed granularity while MPLS paths and ATM circuits have flexible granularity with arbitrary, even zero, bandwidth. The inflexibility in switching granularity can cause serious waste of bandwidth if the granularity of traffic flows does not match that of switching. In a wavelength routed network, the efficiency of multiplexing gain is lost, since each distinct (source, destination) pair require a dedicated channel and traffic demands with fractional wavelength granularity can not be multiplexed together unless they share the same source and destination. Clearly, traffic demands with bandwidth of noninteger wavelength are the rule and those exactly matching integer wavelength are the exception. Apparently, the inadequacy to support the diverse traffic granularity requirement is a serious limitation of wavelength routing.

In addition, the scarcity of wavelengths also limits the scalability of a wavelength-routed network. A circuit in a wavelength-routed network is identified or labeled by the associated wavelength. With current technology, the number of wavelengths is not much greater than 100. In comparison, ATM has 24 bits of VPI + VCI field, and MPLS has 20 bits of path label with thee option to stack labels, each providing millions of route identifiers.

There is an intense effort within IETF to converge the control plane of the optical networks and Internet protocol (IP) networks, which is embodied in the generalized MPLS (GMPLS) proposal [5]. Within the GMPLS framework, wavelength is treated as a label, much the same way as a native MPLS label or ATM VPI/VCI. But interoperability of wavelength paths and MPLS paths or ATM circuits is severely limited by the unique characteristics of wavelength paths, i.e., large and fixed granularity and scarcity of the number of wavelengths.

To overcome the above limitations of wavelength routed optical networks, we proposed a hybrid routing scheme: hybrid wavelength and subwavelength routing (HWSR) [6]. The general idea is to route traffic demands with large granularity using wavelength routing and those with small granularity using subwavelength routing, i.e., performing electronic switching beneath the wavelength level. The benefits of the hybrid scheme is to enable optical bypass for large traffic demands at the same time multiplex multiple small traffic demands into a single wavelength channel to conserve wavelength and better cope

Manuscript received July 30, 2002; revised March 20, 2003.

H. Huang is with the Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, NM 88003 USA (e-mail: hhuang@nmsu.edu).

J. A. Copeland is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30318 USA.

Digital Object Identifier 10.1109/JSAC.2003.815841

with the diverse traffic demand granularity. In this paper, we present both a static optimization formulation and a dynamic heuristic of the hybrid wavelength and subwavelength routing problem.

Similar problem has been studied in the context of SONET wavelength division multiplexing (WDM) ring network traffic grooming [7]–[12], where the problem is how to pack time division multiplexing (TDM) circuits into wavelength channels with the goal of minimizing the number of electronic add–drop multiplexer (ADM). The circuits (traffic demands) are in discrete granularities, e.g., OC-3, OC-12, OC-48, etc. Recently, research on discrete circuit traffic grooming in mesh networks also begins to appear [13], where it is recognized that traffic grooming problem is a variant of network logical topology design problem [16]–[18], with the difference that traditional optical network logical topology design works on the granularity of light paths or wavelength channels and traffic grooming works on finer granularities, i.e., those of digital circuits. The first part of our work deals with traffic grooming in mesh optical networks under static traffic condition, but it is different from [13] in that we consider arbitrary bandwidth traffic flows rather than discrete TDM circuits. Although current optical networks still use SONET style discrete digital hierarchy, i.e., OC-X circuits, future optical networks will predominately carry Internet protocol (IP) traffic, which do not necessarily match SONET style digital circuits that are a legacy of traditional voice networks.

Most of the previous work on traffic grooming is under static traffic condition. Dynamical traffic grooming is now of great interests, because carriers are starting to offer on-demand circuit provision service as a competitive edge. One recent work [15] proposed some heuristics to dynamically groom traffic in mesh networks, using shortest path routing algorithm with the path weight being the number of total or extra light path hops or wavelength links required to establish a connection. The second part of our work also tackles the dynamic traffic grooming problem, but it differs from [15] in that we use the cost of establishing a wavelength channel as criterion, allowing partially filled dedicated wavelength channels if cost justifies so, and our algorithm have explicit load balancing mechanism by incorporating balanced k-shortest path algorithm.

This paper is organized as the following. We introduce the architecture of the hybrid scheme in Section II. We investigate the crucial issue of the hybrid scheme, i.e., the partition and routing of wavelength and subwavelength routed channels under two scenarios: a static optimization formulation and a dynamic heuristic, which are presented in Section III and Section IV, respectively. We conclude the paper in Section V.

II. Architecture of HWSR

In the HWSR scheme, wavelength channels are partitioned into two sets: dedicated and shared channels. Wavelength routing takes place in the set of dedicated channels, each of which is used exclusively by one connection (i.e., a source destination pair) only. Subwavelength routing takes place in the set of shared channels, each of which can be shared by multiple connections. Upon a connection arrival, a decision is made as to whether it should be routed using dedicated or shared wavelength channels based on some policy. If the decision is to use dedicated channels, a wavelength is returned as label, otherwise a (wavelength, sublabel) tuple is returned. The sublabel is an electronic label that enables sharing of wavelength channels, and could take a number of forms such as an MPLS label, an ATM virtual path/circuit index, or a TDM time slot.

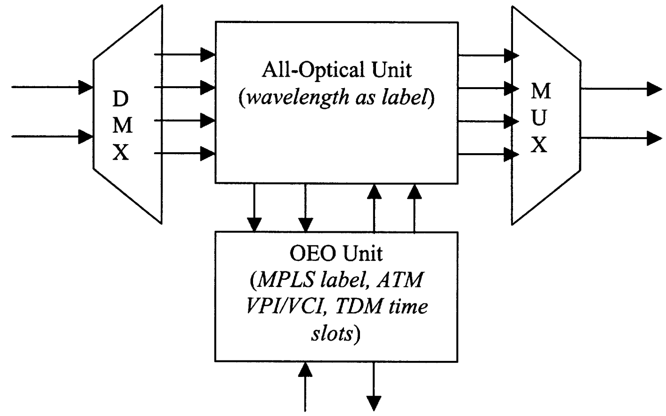

Fig. 1. Structure of an HWSR optical switch.

The node structure of HWSR is shown in Fig. 1. It basically consists of an optical unit and an electrical unit, both of which are standard devices and no novel equipment is required. The optical unit is a regular all-optical switch (wavelength router); it uses wavelength as the switching label and is responsible for implementing dedicated channels. The electrical unit can be an IP router, ATM switch, or SONET DXC and is responsible for implementing shared channels.

Compared with a wavelength-routed network, HWSR has the following benefits. First, multiple connections with fractional wavelength demands can be multiplexed into shared wavelength channels, realizing the benefit of resource sharing and multiplexing gain that is absent in a wavelength-routed network. Second, because of the sharing of wavelength channels, the number of admissible connections in the network is increased, mitigating the problem of scarce wavelength label. However, HWSR needs the extra functionality at nodes for the shared wavelength channels to convert a signal from optical to electrical and again to optical format (OEO conversion), which leads to higher node cost than the wavelength routed counterpart. So, the crucial issue is how to obtain a right balance between dedicated and shared wavelength channels. At one extreme, all wavelength channels are dedicated, which corresponds to the wavelength-routed (WR) or all-optical networks. At the other extreme, all wavelength channels are shared, which corresponds to subwavelength-routed (SWR) or the traditional electronically switched networks. The hybrid scheme HWSR lies in between and offers an optimal point between the relative merits of electrical and optical switching.

III. Static HWSR Routing Optimization

In this section, we are concerned with the situation where traffic demands are known a priori and introduce an integer linear programming (ILP) formulation of the hybrid routing problem. We start with a precise statement of the problem, then introduce our solution approach and the optimization formulation, finally present some numerical results.

A. Problem Statement

The goal of static HWSR design is to satisfy given traffic demands cost-effectively. The design decisions include: how to partition the wavelength channels into dedicated and shared ones, how to construct paths through each set of channels, and how to assign traffic to the paths. The objective is to minimize the system cost of the network and yet still provide adequate resources to support given traffic demands. The system cost consists mainly of two components: node and link costs. Link cost is essentially the cost of wavelength links. Node cost or switching cost depends on the number and type of switch ports utilized. Ports in a nHWSR node can operate in two modes: wavelength routing and subwavelength routing. Subwavelength routing ports are more expensive than wavelength routing ports, since subwavelength routing requires OEO conversion and switching in fine granularity. We differentiate two types of ports by using different port cost weight factors.

We assume traffic demand matrix $\{d_{i,i'}\}$, the amount of traffic requested between the node pair (i, i'), is known *a priori* and it is expressed in the unit of wavelength channels. A noteworthy feature of our traffic demand matrix is that it can take arbitrary fractional values, which is different from previous similar studies that assume traffic matrix of integer or discrete fraction of wavelengths. We also assume that wavelength converters are available at each node, all the demands and the wavelength links are bidirectional and the traffic is routed bidirectionally.

The problem of optimal HWSR network design can be stated formally as the following: given a traffic demand matrix and network topology, find a partition of the wavelength channels into dedicated and shared channels, a sets of routes for each type of channels, and an assignment of traffic to those routes, with the objective of minimizing total system cost. Our problem is similar to optical network virtual topology design, except for the network partition part.

B. Solution Approach

We formulate the HWSR design problem using ILP. Solving ILP is a hard problem, the complexity of which grows exponentially with the size of the problem. The size of the problem depends on the number of the decision variables. In a network of N nodes and E links, each connection has E decisions to make to either include or exclude a certain link, and there could be $N \times (N-1)/2$ connections assuming a fully mesh traffic pattern. Furthermore, each link could be either dedicated or shared, which increases the number of decisions to be made by a factor of two. So the size of the problem is $E \times N \times (N-1)$, i.e., $O(N^3)$ for a sparsely connected network or $O(N^4)$ for a densely connected network, which is a large number even for a network of moderate size.

To make the problem tractable, we precompute a pool of candidate routes. For each connection, i.e., a source and destination pair, we only admit k paths, which are the shortest, the second shortest, the third shortest, ..., and the kth shortest, using Yen's algorithm [14]. By adjusting the value of k, we can vary the complexity of the problem and, thus, the computation time. If k is set to one, we obtain the usual shortest path routing. With larger kk, we obtain a more accurate solution at the expense of computation time. Usually, a moderate value of k, say in the range of 4–8, would suffice to derive a decent solution, since an optimal solution rarely admits long detours. By restriction of our choice of routes, we can reduce the number of route decision variables to the order of $N \times (N-1) \times k/2$, with each connection contributing k route choices. Taking consideration of the partition of dedicated and shared channels, we have a problem size of $N \times (N-1) \times k$, i.e., $O(N^2)$, which is a significant reduction from $O(N^3)$ or $O(N^4)$.

List 1: Notation used in the optimization formulation.

$bw_{j,j'}$	the number of utilized wavelengths on link (j, j′).
wp_i, swp_i	the numbers of utilized wavelength routing and subwavelength routing ports at node i, respectively.
α, β, γ	the unit costs of a wavelength link, a wavelength routing port, and a subwavelength routing port, respectively.
$d_{i,i'}$	the traffic demand of the source and destination pair (i, i′).
$dp^{i,i'}{}_k$	the bandwidth allocated along the kth path for traffic demand $d_{i,i'}$ using dedicated wavelength channels.
$sp^{i,i'}{}_k$	the bandwidth allocated along the kth path for traffic demand $d_{i,i'}$ using dedicated wavelength channels.
$dbw_{jj'}$, $sbw_{jj'}$	the number of dedicated and shared wavelength channels used on link (j, j′).
$W_{j,j'}$	wavelength capacity at link (j, j′).
P_i	switch port capacity at node i.

C. An ILP Formulation

In the following, we present an ILP formulation of the optimal HWSR routing problem. To facilitate the presentation of the formulation, we list the symbols used in List 1. Regarding the switching port, we should clarify that to implement a wavelength routing port, we only need an all-optical port; but to implement a subwavelength routing port, we need one electrical port on the electrical unit and two optical ports going in and out of the optical unit, refer to Fig. 1, and such fact is incorporated into the cost factor γ. The objective function of ILP takes the following form:

$$\sum_{j,j'} \alpha \times bw_{j,j'} + \sum_i \beta \times wp_i + \sum_i \gamma \times swp_i. \quad (1)$$

In the above equation, the first term represents the total link bandwidth cost, and the second and third terms represent wavelength routing and subwavelength routing switch costs, respectively. Here, we have ignored the cost of add–drop ports, which is fixed under static traffic pattern.

Now, we list the set of constrain equations of ILP. For each source and destination pair, a combination of dedicated and shared wavelength paths are used to satisfy demand, as indicated in the following equation:

$$\sum_k dp_k^{i,i'} + \sum_k sp_k^{i,i'} \geq d^{i,i'}. \quad (2)$$

In the above equation, the "larger than or equal to" sign indicates that the amount of bandwidth provided could be larger than the demand. For instance, it might be cost-effective to provide one dedicated wavelength channel if the demand is 0.9 wavelength. The amount of traffic on dedicated wavelength paths must be an integer number of wavelengths, which is embodied in the following constraint:

$$dp_k^{i,i'} \in Z. \quad (3)$$

The wavelength channels on each link (j, j′) are grouped into dedicated and shared ones, the number of each type reflects the sum of the corresponding load from dedicated or shared wavelength paths. Further, the total number of wavelength channels as well as the numbers of dedicated and shared channels used on link (j, j') must all be integers in the unit of wavelength. Thus, we have the following constraints:

$$\sum_{i,i'} \sum_{(j,j') \in p_k^{i,i'}} dp_k^{i,i'} = dbw_{j,j'} \quad (4)$$

$$\sum_{i,i'} \sum_{(j,j') \in p_k^{i,i'}} sp_k^{i,i'} \leq sbw_{j,j'} \quad (5)$$

$$bw_{j,j'} = dbw_{j,j'} + sbw_{j,j'}, \quad (6)$$

$$bw_{j,j'}, dbw_{j,j'}, sbw_{j,j'}, \in Z. \quad (7)$$

Note that the summations in the first two equations are over all source and destination pairs and over all paths passing through the link (j, j'). Note also that ${dp^{i,i'}}_k$ take integer values and occupy dedicated wavelength channels. On the other hand, ${sp^{i,i'}}_k$ may take fractional numbers, but the summation of them in the individual links must be less than the number of shared wavelength channels $(sbw_{i,i'})$, which take integer values and are shared among multiple fractional flows.

The number of utilized wavelength and subwavelength routing ports at node i are calculated in the following:

$$wp_i = \sum_j dbw_{i,j} \quad (8)$$

$$swp_i = \sum_j sbw_{i,j}. \quad (9)$$

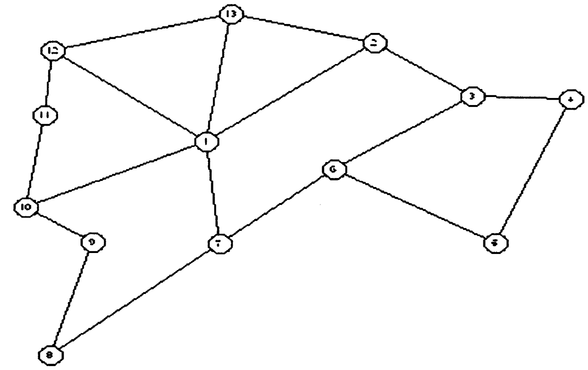

Fig. 2. 13-node 18-link network topology used in the simulation.

Finally, the number of wavelength and switch port must observe capacity constraint.

$$bw_{j,j'} \leq W_{j,j'} \quad (10)$$

$$wp_i + swp_i \leq P_i. \quad (11)$$

With the objective function and constraint equations defined as above, we can solve for the optimal routing of HWSR. The output is a set of dedicated and shared paths that satisfies traffic demand and a partition of switch ports into wavelength and subwavelength routing ones that achieves the minimal total system cost. In the next section, we present some numerical results to examine the performance of HWSR.

D. Numerical Results

The purpose of the numerical study is to provide a quantitative analysis on the system cost improvement of HWSR over those of all-optical switching or WR and electronic switching or subwavelength routing, and how the device cost structure will impact the partition of dedicated and shared channels. Recall from last section, we use α, β, γ to represent unit costs of wavelength links, wavelength and subwavelength routing ports. Different cost scenarios are simulated by varying the values of these parameters. Since the cost relationship between optics and electronics is fluid, changing rapidly when new progress in optical device technology is being made. We choose to study a range of cost scenarios, rather than a fixed one that is necessarily tied to a particular point in the timeline of technology evolution. Because the comparison depends only on the relative values of α, β, γ, we will set α to one in the following without loss of generality.

Our study is based on a realistic network that is provided by Hitachi Telecom, which has 13 nodes and 18 links, as shown in Fig. 2. The number of wavelength per fiber is 80 and the number of fiber per link is four. The number of candidate paths (k) used in our calculation is six. Connections need to be established among about 80% of all node pairs in the network and the traffic demand of each connection is a random real number uniformly distributed in the interval [0, 20]. The ILP problem is solved using a commercial optimization software package CPLEX.

The results shown in Figs. 3 and Fig. 4 reflect the influence of different cost factors on the system cost and the distribution of wavelength channels between dedicated and shared ones. Since cost data revealed by vendors are vague and differ from each

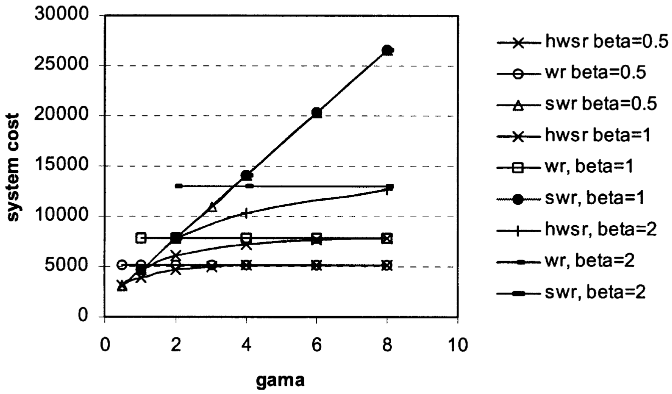

Fig. 3. Total system cost comparison among HWSR, WR, and SWR.

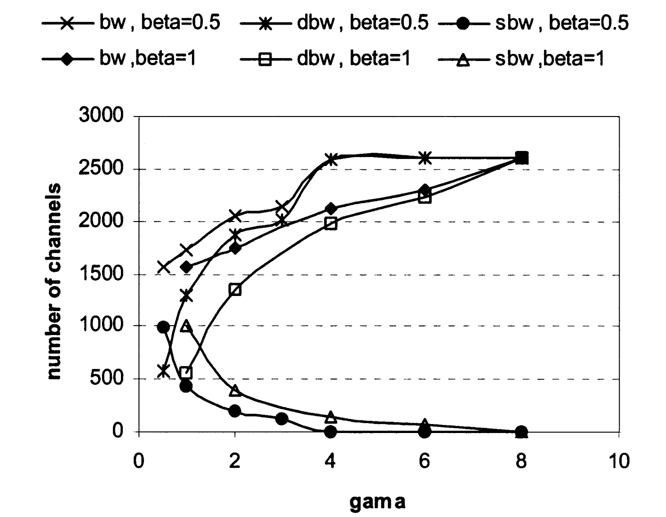

Fig. 4. Numbers of total (bw), dedicated (dbw), and shared (sbw) wavelength channels.

other, in addition device technology is advancing at rapid pace; our choice of cost factors in the following example study does not accurately reflect reality at this moment and only serves for illustrative purpose, but the analysis method does not change and the general features of the plots stay the same, which is verified by our simulations over a wide range of parameters. Fig. 3 shows how the total system costs compare among HWSR, WR, and SWR, with $\beta = 0.5, 1, 2$, and varying values of γ. Note the system cost of WR is independent of the cost of subwavelength routing port (γ), but that of SWR is in linear proportion to γ. From the figure, it is evident that HWSR always incurs the least cost, which is not surprising since WR and SWR are just two extreme cases of HWSR. It also can be seen from the figure that HWSR is most effective when the tension between optical bypass and multiplexing gain is most intense, which is around the region where the corresponding curves of WR and SWR cross.

Fig. 4 shows how the numbers of dedicated and shared wavelength channels vary with γ, with β again set to 0.5 and 1. As expected, the number of shared channels decreases when the cost of subwavelength port increases, but the total number of wavelength channels increases with increasing γ, which is due to the loss of resource sharing or multiplexing gain with increasing cost of doing subwavelength routing.

For other β values, we have very similar figures as Figs. 3 and 4, which are omitted here because of redundancy. We also performed simulation on a larger network with 21 nodes and 29 links. The results are very similar and again are omitted here because of lack of noteworthy features.

IV. Dynamic HWSR Routing Heuristic

In the previous section, we attacked hybrid wavelength and subwavelength routing problem using an optimization formulation under the static traffic condition. In practicality, traffic demand is hard to predict because of the fluidity of market size, competition posture, and general economic situation. Therefore, a dynamic HWSR scheme that does not known the traffic matrix a priori and makes routing decision upon each connection arrival is highly desirable. In this section, we first present a dynamic traffic model, then introduce a dynamic routing heuristic and an analysis to determine its parameters, and finally present some numerical results and compare them to those obtained using static optimization.

A. Traffic Model: Incremental Arrival With Sporadic Random Termination (IASRT)

We consider the situation where traffic requests arrive dynamically. Currently, optical networks mainly serve as backbone networks and traffic connections are essentially trunks, which have different characteristics than those of so-called microflows, such as a telephone conversation or a transmission control protocol (TCP) session. Trunks usually represent circuits that are sold or leased by wholesale service providers to retail service providers or large business customers. They normally have much longer lifetime than microflows; once set up, seldom get rerouted and rarely torn down unless customers cancel orders or the network needs major upgrade. Typically, an optical network starts out as empty. Then, there is a rise period when many new connections are created rapidly with few old connection terminations in between. After the rise period, the network arrives at a stable state where the majority of connections have been and remains established with a small number of connection arrivals and departures. The vast opportunity to do optimization is in the rise period when the network's overall route configuration takes shape. Once reaching the stable state, the network is fully shaped and further optimization faces many restrictions since rerouting existing routes is generally not advisable. In addition, stable state may be brief since it indicates the network's capacity is fully utilized and a major update is imminent. Therefore, our main concern is with the rise period, where Poisson modeling is clearly not adequate. To model the rise period, we assume an IASRT, where new connections are established one by one with a few connection termination event randomly interspersed until the capacity of the network is nearly exhausted.

Accordingly, the metric of performance merit needs a rethinking. With IASRT, we are less interested in blocking probability as in a typical Poisson modeling; rather we are concerned with minimizing transport cost to support certain traffic demands, which means either less amount of network resource is used to support given traffic demands or more traffic demands can be supported using the same amount of resource.

B. Dynamic Hybrid Routing Heuristic

The design goal of the dynamic hybrid routing heuristic is to emulate the effect of the static optimization of the previous section, which achieves minimum cost because its complete traffic knowledge. To such end, we need to route traffic using minimal network resources, which means shorter paths are always preferred except where load balance is needed. We also need to partition traffic judiciously into dedicated and shared channels so that integer or near integer wavelength flows go to dedicated channels and small fractional flows go to shared channels. For the path routing part, we used a balanced k-shortest path algorithm, listed in List 2, which modifies the k-shortest path algorithm [14] by randomize the path returned when multiple paths have the same length to induce load balance effect.

List 2: Balanced k-shortest path algorithm.

A. For each source and destination pair, compute a set of k-shortest paths, put the set into a queue Q1, rank each path with a score according to its length, with higher score having shorter length.
B. If two or more paths have equal length, they have the same score but their relative position in Q1 is initialized randomly.
C. For each call to the algorithm, return Q1 and,
D. Randomly reshuffle the segments of Q1 with equal scores.

For the traffic partition part, we adopt the policy shown in List 3.

List 3: Traffic partition policy.

$$b^i \begin{cases} \lfloor b^i \rfloor \longrightarrow dedicated\ \ channel \\ b^i - \lfloor b^i \rfloor \begin{cases} if >= p \longrightarrow dedicated\ \ channel \\ if < p \longrightarrow \ \ shared\ \ channel. \end{cases} \end{cases}$$

The traffic partition policy always route integer wavelength portion of a connection using dedicated channels. The fractional wavelength portion is also routed using dedicated channels if it is larger than a threshold p; otherwise, it is routed using shared channels. Again, a dedicated channel can be used by only one connection, but a shared connection can by used by many connections until its capacity is exhausted. The determination of value p will be discussed in the next section.

List 4: Dynamic hybrid routing heuristics.
For each connection b^i do the following.

A. Partition b^i into an integer wavelength part and a fractional wavelength part.
B. Select a path with highest score from the k-shortest path set that has enough port and link capacity to route the integer wavelength portion of b^i using dedicated channels. If a path is found, return the path and update wavelength routing port count and link capacities; otherwise reject the connection.
C. For the fractional portion of b^i, use traffic partition policy to determine if dedicated or shared channels should be used. If dedicated channels are to be used, go to B except using the fractional wavelength part of b^i. If shared channels are to be used, go to D.
D. Select a path with highest score from the k-shortest path set that has enough port and link capacity to route the fractional wavelength portion of b^i using shared channels. If a path is found, return the path and update subwavelength routing port count and link capacities; otherwise reject the connection.

Synthesizing the above two parts, we have a dynamic hybrid routing heuristic as shown in List 4.

C. Determination of Traffic Partition Threshold

List 5: Notation used in the dynamic hybrid routing heuristic.

b^i	the amount of bandwidth requested by ith connection.
f^i	the amount of fractional bandwidth of the ith connection.
${C^i}_d, {C^i}_s$	the transport costs for the ith connection using WR and SWR. respectively
h^i	the number of hops for ith connection.
p	traffic partition threshold parameter.

(Amount of bandwidth are expressed in the unit of wavelength)

A central concern in our determination of the threshold parameter is to minimize transport cost to support a connection. The transport cost of a connection is determined by the connection's resource consumption, which depends on wavelength mileage, and the number and type of switch ports consumed by the connection, scaled by the bandwidth of the connection and weighted by appropriate factors. Again, in order to facilitate our discussion we list notations used in List 5, in addition to those introduced in the previous section. Generally, the transport cost of a connection with bandwidth b^i and hop count h^i can be expressed as below, again using the fact that each hop is terminated by two switch ports and add–drop ports represent a fixed cost and, therefore, its effect ignored

$$C^i = (\alpha + 2(\beta \; or \; \gamma))\, h^i b^i.$$

Now, we specialize to connections using dedicated and shared wavelength channels. In the case of a connection using dedicated channels, the bandwidth request could be an arbitrary amount but the allocated channels must be integer number of wavelengths. So a packing overhead is introduced using WR to satisfy a connection i with bandwidth request (b^i); the transport cost is given as follows:

$$C^i{}_d = (\alpha + 2\beta)h^i[b^i]$$

where the operator $[\cdot]$ indicates an integer ceiling operation. On the other hand, if a connection with fractional wavelength demand (f^i) uses shared channels, the unused bandwidth can be utilized by other connections; consequently, there is no packing overhead and the transport cost is given as

$$C^i{}_s = (\alpha + 2\gamma)h^i f^i.$$

Summing over all connections using either dedicated or shared channels, we can get total transport cost.

Now, we turn to the determination of p. The fractional wavelength part f^i of a connection b^i can use either dedicated or shared channels depending which one has lower cost. If dedicated channels are used, the transport cost is given as follows using the fact the $[f^i] = 1$:

$$C^i{}_d = (\alpha + 2\beta)h^i[f^i] = (\alpha + 2\beta)h^i.$$

On the other hand, if shared channels are also used, the transport cost is given as

$$C_s{}^i = (\alpha + 2\beta)h^i f^i.$$

At threshold condition, i.e., when fractional part f^i is equal to the threshold value p^i, the two costs are equal, and we have

$$(\alpha + 2\beta)h^i = (\alpha + 2\gamma)h^i p^i$$

or

$$p = \frac{\alpha + 2\beta}{\alpha + 2\gamma}$$

which means the threshold is equal to the ratio of costs to implement a dedicated and a shared channel, recall that each channel consists of one link and two ports. The result is surprisingly simple but intuitively correct, since if a wavelength fraction is larger than the cost ratio, compacting overhead does not matter anymore and it is always cost effective to use dedicated channels.

D. Simulation Results

The purpose of the simulation is to study the performance of the dynamic heuristic by comparing it with the static optimization. To that end, we used the same network topology, link capacity and port number, and the same partition between wavelength routing and subwavelength routing ports. Furthermore, we use the same traffic matrix as the static optimization and serialize it for dynamic heuristic consumption. In the static case, each traffic matrix element, i.e., traffic demand between a certain source and destination pair, has on average ten wavelength demand. In the dynamic case, we randomly partition each static traffic matrix element into ten incremental demands and randomly intermingle these incremental demands with those from other traffic matrix elements to form a serialized incremental traffic arrival queue Q2. We simulate the incremental part of IASRT by dequeueing Q2, and the random termination part by randomly terminating an old flow every ten new arrivals but reinserting the old flow back to the end of Q2 to conserve traffic matrix equality between the static and dynamic cases.

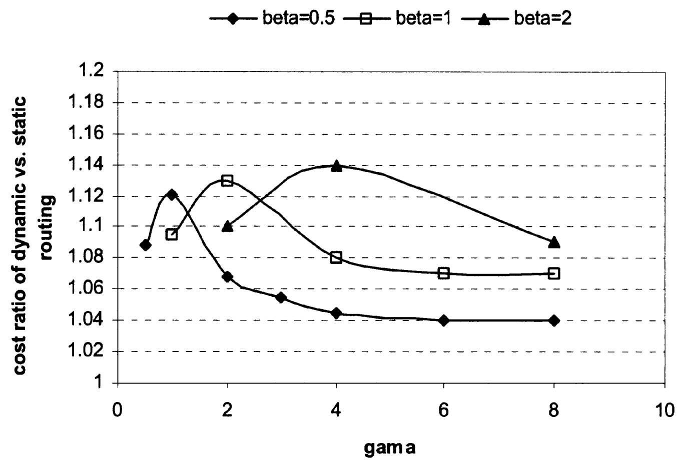

Fig. 5. Transport cost ratio of the dynamic heuristics versus those of the static optimization.

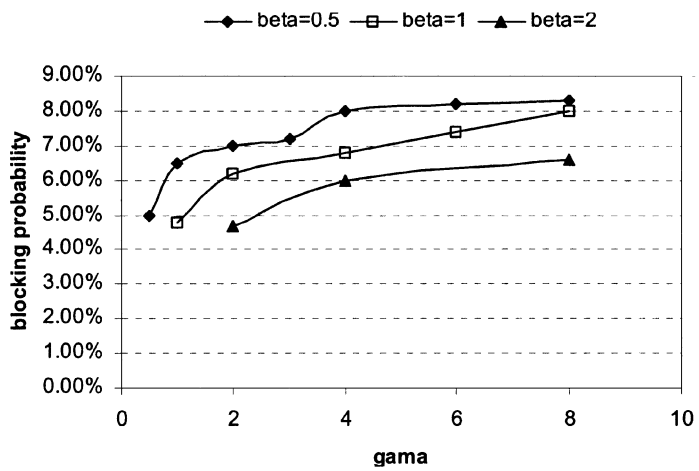

Fig. 6. Blocking probabilities of the dynamic heuristic.

Our results are summarized in Figs. 5 and 6. Recall from List 1, we use α, β, γ to represent unit costs of wavelength links, wavelength and subwavelength routing ports. In the figures, again β is set to 0.5, 1, 2, with γ varied. Other values of β are also simulated but they produce similar results and, thus, omitted here. The total transport costs incurred using dynamic heuristic versus that of static optimization are plotted in Fig. 5. Apparently, the dynamic heuristic tracks the static optimization performance quite well, resulting in no more than

14% cost increase in the example network. The peaks in the relative cost curves correspond to regions where HWSR is neither closer to WR nor to SWR, so choice of channel partition becomes more important and incurs more errors under dynamic traffic than static one. Blocking probabilities of different β and γ combinations for the dynamic heuristic are plotted in Fig. 6. There is no blocking in the static optimization; however a small blocking (no more than 9% in our case) occurs in the dynamic heuristic, because it makes decisions one connection arrival at a time without exploiting global traffic knowledge and some connections that are otherwise routable in a static optimization get rejected in the dynamic heuristic. It is not surprising that blocking is more severe in the region where channel sharing is expensive, i.e., large γ.

V. Conclusion

In this paper, we proposed a hybrid wavelength and subwavelength architecture, presented a static optimization formulation and a dynamic heuristics. The hybrid routing scheme retains the benefits of optical bypass at the same time provides bandwidth diversity and resource sharing that is inherent in Internet traffic and IP centric schemes such as MPLS. We hope our work can stimulate further thinking on converging optical and IP networking.

References

[1] B. Mukherjee, *Optical Communication Networks*. New York: McGraw-Hill, 1997.

[2] R. Ramaswami and K. Sivarajan, *Optical Networks*. San Mateo, CA: Morgan Kaufmann, 1998.

[3] P. Green, "Progress in optical networking," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 54–61, 2001.

[4] B. Mukherjee, "WDM optical communication networks: Progress and challenges," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1810–1824, Oct. 2000.

[5] P. Ashwood-Smith *et al.*, "Generalized multi-protocol label switching (GMPLS) architecture," *IETF Internet Draft*, Mar. 2002.

[6] H. Huang and J. A. Copeland, "Hybrid wavelength and sub-wavelength routed optical networks," in *Proc. IEEE GLOBECOM*, Nov. 2001, pp. 2119–2123.

[7] O. Gerstel, R. Ramaswami, and G. Sasaki, "Cost-effective traffic grooming in WDR rings," in *Proc. IEEE INFOCOM*, Mar.–Apr. 1998, pp. 69–77.

[8] A. L. Chiu and E. H. Modiano, "Reducing electronic multiplexing costs in unidirectional SONET/WDM ring networks via efficient traffic grooming," in *Proc. IEEE GLOBECOM*, Nov. 1998, pp. 322–327.

[9] X. Zhang and C. Qiao, "An effective and comprehensive solution to traffic grooming and wavelength assignment in SONET/WDM rings," *Proc. SPIE*, vol. 3531, pp. 221–232, 1998.

[10] O. Gerstel, P. Lin, and G. Sasaki, "Combined WDM and SONET network design," in *Proc. IEEE INFOCOM*, Mar. 1999, pp. 734–743.

[11] W. Cho, J. Wang, B. Mukherjee, B. Moslehi, and R. J. Black, "Unification and extension of the traffic-grooming problem in WDM ring networks," in *Proc. Optical Fiber Communication Conf. (OFC)*, Mar. 2000, pp. 162–164.

[12] J. Wang, W. Cho, V. R. Vermuri, and B. Mukherjee, "Improved approaches for cost-effective traffic grooming in WDM ring networks," in *Proc. IEEE Int. Conf. Communications (ICC)*, June 2000, pp. 1295–1299.

[13] K. Zhu and B. Mukherjee, "Traffic grooming in an optical WDM mesh network," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 122–133, Jan. 2002.

[14] J. Y. Yen, "Finding k shortest loopless paths in a network," *Manage. Sci.*, vol. 17, no. 11, pp. 712–716, July 1971.

[15] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "Dynamic traffic grooming in WDM mesh networks using a novel graph model," in *Proc. IEEE GLOBECOM*, Nov. 2002, pp. 2681–2685.

[16] R. Ramaswami and K. N. Sivarajan, "Design of logical topologies for wavelength-routed optical networks," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 840–851, June 1996.

[17] D. Banerjee and B. Mukherjee, "Wavelength-routed optical networks: Linear formulation, resource budgeting tradeoffs, and a reconfiguration study," *IEEE/ACM Trans. Networking*, vol. 8, pp. 598–607, Oct. 2000.

[18] E. Leonardi, M. Mellia, and M. A. Marsan, "Algorithm for the logical topology design in WDM all-optical networks," *Opt. Networks Mag.*, vol. 1, no. 1, pp. 35–46, Jan. 2000.

Hong Huang (S'99–M'02) received the B.Engr. degree from Tsinghua University, Beijing, China, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta.

Currently, he is an Assistant Professor in the Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces. His research interests include optical networks, wireless networks, MPLS, and IP quality-of-service.

John A. Copeland (M'68–SM'76–F'83) received the Ph.D. degree from the Georgia Institute of Technology, Atlanta.

He is currently John H. Weitnauer, Jr. Technology Transfer Chair, GRA Eminent Scholar, Director of the Communication Systems Center, and Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology. His research interests include computer networks, digital CATV networks, computer architecture and operating systems, and optical networks.

Dr. Copeland is a recipient of the IEEE Morris N. Liebmann Award, served as Editor of the IEEE Transactions on Electron Devices, and was on the Board of Trustees of the Georgia Tech Research Corporation.