Routing of Multicast Audio/Video Streams in Reconfigurable WDM Optical Networks*

Ciro A. Noronha Jr.† Fouad A. Tobagi

Department of Electrical Engineering
ERL 454, MC 4055
Stanford University
Stanford, CA 94305-4055
email: {ciro,tobagi}@bodega.stanford.edu
Phone: +1 (415) 723-1708
Fax: + 1 (415) 725-7398

Abstract

Due to its low attenuation, fiber has become the medium of choice for point-to-point links. Using Wavelength-Division Multiplexing (WDM), many independent channels can be created in the same fiber. A network node equipped with a tunable optical transmitter can select any of these channels for sending data. An optical interconnection combines the signal from the various transmitters in the network, and makes it available to the optical receivers, which may also be tunable. By properly tuning transmitters and/or receivers, point-to-point links can be dynamically created and destroyed. Therefore, in a WDM network, the routing algorithm has an additional degree of freedom compared to traditional networks: it can modify the network topology to create the routes. In this paper, we consider the problem of routing multicast audio/video streams in WDM networks and propose heuristic algorithms to solve it. The performance of these heuristics is evaluated in a number of scenarios, with a realistic traffic model, and from the evaluation we derive guidelines for usage of the proposed algorithms.

---

*This work was supported in part by NASA under grant NAG2-842, by the National Science Foundation under grant NCR-9016032 and by Pacific Bell. Ciro Noronha was supported by a graduate scholarship from FAPESP from Sept/89 to Aug/93 under grant 89/1658.

†Currently with Optivision, Inc., Palo Alto, CA 94304

1 Introduction

Due to its low attenuation (less than 0.2 dB/km) and very high bandwidth, fiber has become the medium of choice for point-to-point links at high speeds, for any distance over $\approx$ 100 m and for data rates of 45 Mb/s and up. Using Wavelength-Division Multiplexing (WDM), many channels can be created in the same fiber. A network node equipped with a tunable transmitter can select any of these channels for sending data, and can dynamically change this selection. The range of wavelengths addressable depends on the technology used to implement the tunable transmitter [1]. The optical signal from the transmitters in the network is combined by an optical interconnection (e.g, a WDM star coupler), and made available to a subset of the optical receivers, determined by the optical interconnection. Optical receivers can also be tunable [1]. The interconnection pattern between nodes is defined by the tuning of transmitters and receivers to specific wavelengths, and can be dynamically changed. In a traditional (fixed-topology) network, given the traffic requirements, the routing algorithm is responsible for finding routes, which must satisfy these requirements. In a WDM network with tunable transmitters and receivers, the routing algorithm has an additional degree of freedom: it can choose (or modify) the topology to *create* the routes as needed.

We assume that the applications generate traffic in *sessions*. A session is a group of video/audio streams that are logically related. For example, a video-conference with P participants where each of the conferees can see all the other participants represents a session with P multipoint streams, from each conferee to the other $P-1$ participants. A *stream* is defined as a continuous flow of information (i.e., video frames or audio samples) that has to be delivered in a timely fashion. Multimedia streams have the following new requirements (when compared to traditional data traffic), which must be taken into account when they are routed:

Bandwidth - multimedia streams use relatively high bandwidth on a continuous basis for long periods of time. For example, a high-quality compressed video stream can use anywhere from 1.5 to 8 Mb/s for extended periods of time, while the average bandwidth used by typical data applications can be well below 1 Mb/s.

Multipoint Communications - it is expected that a significant fraction of the multimedia traffic will be multipoint. Examples are videoconferencing, one-way video distribution and collaborative computing. Data applications, on the other hand, typically make only occasional use of multicasting.

Low Latency (on the order of 100-200 ms end-to-end), required for some applications (such as videoconferencing or collaborative computing) that provide interactive communications. Data applications typically do not have such strict latency constraints.

In this paper, we consider the problem of routing multimedia streams in a WDM network. In section 2, we discuss the characteristics of the optical components used in building a network, and describe the previous work in optical WDM networks. In section 3, we give the problem formulation, and in section 4 we propose a number of heuristic algorithms for multicast routing in the WDM network. Evaluation of the heuristic algorithms is presented in section 5, where we compare the various heuristics proposed under a dynamic traffic environment. Discussion and conclusions are presented in section 6.

2 Optical Network Components and Configurations

In this section, we give a brief overview of the optical network components considered in this paper, as well as of the previous work on optical WDM networks.

2.1 Optical Network Components

The optical WDM network has the following basic "building blocks" [1, 2, 3]:

The Optical Interconnection: The optical interconnection is responsible for mixing the light from the transmitters and splitting (dividing) it among the receivers, irrespective of the wavelength. The most common optical interconnection is the WDM star. The

WDM star equally divides the optical power from each of the incoming ports among the output ports. The optical signal in each output port is a combination of the optical signals from each of the input ports. Other optical interconnections, such as trees, multiple stars, etc., are possible; they do not necessarily divide the input power equally among all the outputs. In general, a given output will receive the signal from a subset of the inputs [4, 5]. In this paper, we assume that the optical interconnection is such that the light signal from all input ports is available to all output ports.

The Optical Transmitters: Optical transmitters are responsible for modulating the optical signal with the user data. Tunable optical transmitters are built using tunable lasers. There are several methods available for building such devices [6, 7]; the different methods represent a tradeoff between tuning speed and tuning range. In this paper, we consider both fixed-wavelength and tunable optical transmitters. The tuning time is considered to be in the sub-millisecond range, which is well within today's technology. Moreover, we assume that the tunable transmitters are able to address all the wavelengths used in the network.

The Optical Receivers: Optical receivers are responsible for extracting the user data from the optical signal. There are basically two types of WDM optical receivers: (i) direct detection receivers, where the desired wavelength is extracted by an optical filter prior to detection [8], and (ii) coherent receivers, where the optical signal is mixed with the light from a local laser (which is kept synchronized with the transmitter laser by means of an optical PLL). Tunable optical receivers can be made by using either tunable optical filters [9] or tunable lasers. In this paper, we assume both fixed-wavelength and tunable optical receivers. The same assumptions made for the tunable optical transmitters on tuning range and speed are repeated for the tunable optical receivers.

2.2 Optical Network Configurations and Operation

We classify the work in WDM networks into two categories: *Single-Hop Networks* and *Multi-Hop Networks*. The work in Single-Hop Networks is characterized by the fact that a direct channel is established between the sender and the receiver, and communication is single-hop. The network interconnection is usually assumed to be a WDM star. In Multi-Hop Networks, the sender and the receiver do not necessarily share a common channel; the message may go from node to node until it reaches the destination. More general network topologies are sometimes considered, but in most cases the the physical interconnection is still the WDM star or some distribution network (such as a tree) that makes the signal from all the transmitters available to all the receivers. A comprehensive review of the work in the field of WDM networks can be found in [10] and [11]. We will not discuss further the Single-Hop Networks since they do not relate to the work presented in this paper.

In [12, 13, 14, 15], a multi-hop network is considered; it is assumed that the network reconfiguration process will be performed infrequently; during the reconfiguration, the network may even be non operational. The problem then becomes similar to a traditional topological design problem, where the traffic is Poisson and the traffic matrix is known, with additional constraints introduced by the fact that each node has a well-defined number of transmitters and receivers. They all consider that the combined optical signal from all transmitters is available to all receivers. The differences are in the following areas:

(i) Objective function to be optimized: in [12], the objective function is to minimize the average delay; the authors assume a queue model for the nodes, which makes the delay a non-linear function of the flow in the links. They also take into account the propagation delays in the network. In [13], the network is assumed to operate under deflection routing, and the average delay is indirectly minimized by minimizing the length of the alternate paths between sources and destinations. In [14, 15], the objective function is to minimize the maximum flow over all links.

(ii) Additional constraints in the optimization: in [15] tunability restrictions are assumed, i.e., receivers can only be tuned to a subset of the available bandwidths. The other papers do not have additional constraints.

(iii) Solution method: in [12, 13] the objective function is non-linear, and the authors use the simulated annealing method to search for a sub-optimal solution. In [14, 15] the authors present a heuristic algorithm which divides the problem into two subproblems - the wavelength assignment subproblem and the routing subproblem, which are solved by linear programming.

In summary, the previous work in the field of WDM networks has assumed that tuning of the transmitters and receivers happens over a very long time scale; the topology of the network does not change often, and when it does, it is in response to changes in the traffic matrix. The problem then becomes similar to the traditional topological design problem, with some additional constraints (i.e., the number of links leading to a node must be equal to the number of receivers in that node, and similarly for transmitters). Reconfiguring the network only when the long-term traffic trends change does not make use of the full switching potential of the WDM network. When dealing with streams, it is possible to reconfigure the network when requests arrive, and when streams terminate. Conceptually, this is similar to the long-term reconfiguration, but the change in traffic trends are actual stream arrivals and terminations. A similar approach is described in [16], but there the authors consider packet bursts instead of streams.

A final observation is that none of the previous work has dealt with multicast (point-to-multipoint) traffic; in all cases, the traffic is considered to be unicast (point-to-point).

3 Problem Formulation

In this section, we define the problem of routing streams in a WDM network. We describe the traffic model, and list the assumptions we make about the optical network. We then define the cost and delay measures, and give the statement of the problem.

3.1 The Traffic Model

For the traffic, we assume that user's requests come in *sessions*. A session is a group of streams that are logically related. We will denote by T the number of streams in the session. Stream i, $i = 1, \ldots, T$, is characterized by its source s_i, its n_i destinations $d_{i1}, d_{i2}, \ldots, d_{in_i}$, its bandwidth requirement r_i and its maximum latency constraint D_i. We assume that all the streams in the session arrive and depart simultaneously, the session arrival process is a Poisson process with rate λ, and the session duration is exponentially distributed with rate μ.

3.2 Optical Network Assumptions

We assume that the network operates in a store-and-forward, multi-hop operation, but the reconfiguration of the network happens in a stream-by-stream basis, creating paths as the streams arrive, and removing them after they terminate. Bandwidth in the links can be shared in a TDM fashion.

We make the following assumptions about the network, which is depicted in Figure 1:

- There are N nodes in the network; node i, $i = 1, \ldots, N$ is equipped with S_i optical transmitters and P_i optical receivers.

- The optical interconnection is such that all receivers have access to the light signal from all transmitters. No other assumptions are made about it. This assumption simplifies the formulation of the problem, but limits the results to the LAN/MAN environment.

- The number of distinct wavelengths, denoted by W, is larger than the number of transmitters/receivers in the network. Due to the large bandwidth available in the fiber, this is a reasonable assumption.

- At any time, only one transmitter can be tuned to a given wavelength. We do not consider multiple-access operation (i.e., many transmitters tuned to the same wavelength), because this is difficult to implement efficiently in optics. On the other hand, if receivers are tunable, multiple receivers can be tuned to a the wavelength used by a given transmitter, providing physical multicasting.

- Any given transmitter can be connected to any given receiver - there are no tunability restrictions. Current technology allows the implementation of transmitters and receivers that are tunable over relatively wide ranges, making this a reasonable assumption.

3.3 The Cost and Delay Measures

The route for a multicast stream is a tree, formed by merging the individual paths from the source node to each of the destinations. To each link in the tree we associate a *cost* and a *delay*. The link is formed by a transmitter/receiver pair, tuned to the same wavelength. The link delay is the sum of the transmitter delay and the receiver delay. For example, assuming a star topology, the transmitter delay will be the propagation delay from the transmitter to the star, and the receiver delay will be the delay from the star to the receiver. We assume that delays are independent of the data flow in the link, which is a reasonable assumption for multimedia traffic [17, 18]. Similarly, the cost of the link is assumed to be the sum of a transmitter cost and a receiver cost.

We define the *multicast cost* as the ***sum*** of the costs of the links in the tree, scaled by the bandwidth used. The delay of a path between two nodes is defined as the sum of the delays of the links in the path. Since interactivity requirements limit the *maximum* delay between two nodes, we define the *multicast delay* as the ***maximum*** of the delays of the paths between the source and all the destinations.

While the physical meaning of the multicast delay is clear, it is important to discuss the physical meaning of the multicast cost, as it is a function of the values assigned to the link costs. Link costs may represent actual monetary costs - e.g., to send a stream of bandwidth b bits/s over a link which costs c \$/bit will cost cb \$/bit. In this case, the multicast cost is the monetary cost of running that stream. Another possible meaning is cost as a measure of the usage of network resources. The simplest case is when all the link costs are set to 1; the multicast cost will be the amount of network bandwidth used in routing the multicast stream. This measure can be refined to reflect the desirability of using a certain link; for example, if a network is composed of two densely connected subgraphs connected by a few links, these few links would have higher costs, to force the routing algorithm to avoid them unless the source and the destination lie in different subgraphs. In an optical network where nodes might have different numbers of transmitters and receivers, costs may be allocated in inverse proportion to their number (i.e., a transmitter on a node that has a large number of transmitters may be "cheap", but another in a node that has just a few may be "expensive"). In this sense, the cost may become a control parameter to be set by the network designer/adminstrator to improve performance. As a matter of fact, this is how link labels are assigned in current networks.

3.4 Statement of the Problem

The problem under consideration can be stated as: "Given a session with T streams, each stream being characterized by its source, destinations, latency constraint and bandwidth requirement, find the logical network topology and routes that satisfy the stream requirements, while optimizing a given objective function."

A solution to the reconfiguration/routing process is composed of two parts: (i) the wavelength assignment, which defines which transmitters are connected (tuned) to which receivers (and thus defines the network topology), and (ii) the assignment of routes given the wavelength assignment. If multiple solutions exist for a reconfiguration/routing problem, the objective function is the criterion used to select the "best" one. Possible objective functions include minimum cost, minimum delay, or a combination of these.

4 Heuristic Algorithms for the Reconfiguration and Routing Problem

The reconfiguration and routing problem, as formulated in section 3, is NP-complete. In [3], we show that the problem can be solved exactly using integer linear programming. However, the optimum algorithm given there has (in the worst case) exponential run time. In this section, we propose a number of simpler heuristic solutions. In [19] we proposed a heuristic algorithm for the unicast case; in this paper, we use that algorithm to build minimum-cost and minimum-delay heuristics for the multicast case.

We present here a summary of the heuristic algorithm for routing unicast streams; a detailed presentation can be found in [3, 19]. Given a session with one or more streams, we seek to find the logical network topology and the routes for this session. From a high level point of view, the algorithm proposed here starts with an arbitrary initial logical topology, and makes changes to it considering the streams in the session one at a time. The changes are made using the *Shortest Path with Reconfiguration Algorithm*, a variation of Dijkstra's Shortest Path algorithm [20] proposed by us that works in a reconfigurable network environment. In the following, we first describe the Shortest Path with Reconfiguration Algorithm, and then give the complete reconfiguration and routing heuristics.

4.1 The Shortest Path with Reconfiguration Algorithm

Our objective is to compute the shortest path in a WDM network, where the topology of the network is a free variable that can also be used to minimize the path length. In the best case, we would just tune a transmitter at the source and a receiver at the destination to the same wavelength, and obtain the shortest possible path, with length equals to one hop. However, this might not always be possible, since transmitters and receivers might be already connected to other nodes. In general, we classify the transmitters and receivers in the network either as *free* or *locked.* Free transmitters and receivers are idle and can be reconfigured at will; locked transmitters and receivers have been tuned and are currently being used by at least one stream. The shortest path algorithm can only reconfigure the free transmitters and receivers, although it might make use of the locked ones in whatever topology they happen to be. If the WDM network supports physical multicasting (i.e., if it has tunable receivers), this is taken into account in the algorithm by *considering all the transmitters as free, without regard to the other connections.* The algorithm described below does exactly this: given the WDM network in a certain logical topology, where some links are free and some are locked, and a source-destination pair, it finds the shortest path between these two nodes, reconfiguring the free links if necessary.

Step 1: Using Dijkstra's algorithm, identify: (i) the shortest path between the source and the destination, and (ii) the node closest to the source which has a free transmitter (i.e., either the source itself or the first node added to the shortest path tree that has a free transmitter); this node, if found, will be denoted by *Node A*. Note that if the network is disconnected, there might not be a path between the source and the destination.

Step 2: Using Dijkstra' algorithm in reverse from the destination to the source (i.e., building the tree in reverse), find the node closest to the destination that has a free receiver; this node, if found, will be denoted by *Node B*.

Step 3: If either node A or node B or both were not found, stop. If a path between the source and the destination was found in step 1, it is the shortest path. Otherwise, there is no path. If both node A and node B were found, proceed to step 4.

Step 4: Let L_1 denote the length of the shortest path found in step 1 (make $L_1 = \infty$ if no path was found), and L_2 denote the length of the path obtained by tuning the transmitter in node A to the receiver in node B, and using the shortest path from the source to A, the newly-created A-B link, and the path from B to the destination. If $L_1 \leq L_2$, do not reconfigure the network and use the shortest path from step 1; otherwise, tune A to B and use the path just created, as described above.

In [16] another algorithm for computing paths in a reconfigurable WDM network is given. If the metric is just number of hops, the algorithm given there is similar to the shortest path with reconfiguration presented here, with the difference that the routes there are computed by exhaustive search rather than by using shortest path.

4.2 The Reconfiguration and Routing Heuristic for Unicast Streams

Given a session, the basic idea behind the wavelength assignment heuristic is to take an arbitrary initial topology, and apply the shortest path with reconfiguration to each of the components of the session. The shortest path with reconfiguration algorithm is applied to the streams in the session in decreasing order of bandwidth. In general, the first streams to be routed will be given shorter paths, as more network resources are available. Therefore, it is better to route first the higher-bandwidth streams, to minimize their usage of network resources.

Step 1: Choose an arbitrary initial wavelength assignment. Create a vector $\boldsymbol{U}$, containing the used bandwidth on each transmitter; initially, $U_i = 0$, $i = 1, \ldots, K$. Sort the streams in the session in order of bandwidth.

Step 2: Consider the stream with the highest bandwidth requirement that was not yet processed; let us denote it by stream j. Temporarily prune from the network topology the transmitter/receiver pairs that do not have enough free bandwidth to support the stream, i.e., belonging to the set $\{i : V - U_i < r_j\}$, where V is the link bandwidth. Mark all the transmitter/receiver pairs belonging to the set $\{i : U_i > 0\}$ as locked, and the remainder as free.

Step 3: Execute the "Shortest Path with Reconfiguration Algorithm" described earlier for this stream. If successful, update the $\boldsymbol{U}$ vector as follows: $U_i \leftarrow U_i - r_j$, $i \in$ path.

Step 4: If all streams in the session have been considered, terminate; otherwise, return to step 2.

After this algorithm is run, the initial network topology is transformed into a new topology which matches the session requirements. If all the invocations of the shortest path with reconfiguration algorithm in step 3 are successful, a set of routes for the session is also available; otherwise, the heuristic fails.

4.3 Heuristic Algorithms for Multicast Routing In WDM Networks

In this section, we present minimum-cost and minimum-delay heuristic algorithms for WDM networks, based on the Shortest Path with Reconfiguration algorithm. The algorithms presented here are for a single multicast stream; bandwidth constraints are taken care of by pruning from the network those links with insufficient free capacity, and multiple streams are handled sequentially, as described in the previous section.

4.3.1 Minimum Delay Heuristic Algorithm

On a fixed-topology network, minimum-delay multicast routing can be achieved by finding the shortest path from the source to each of the destinations (using the link delays as labels) and then merging these paths. For a single multicast, this is an exact algorithm, i.e., the routes found are optimal. The same idea can be applied to the WDM network: one would use the Shortest Path with Reconfiguration algorithm between the source and each of the destinations, and merge the paths. However, unlike its fixed-topology counterpart, this algorithm might not be optimal even for a single multicast stream, depending on whether or not physical multicasting is available.

In the fixed-topology network, using shortest path leads to minimum-delay paths. If this process is repeated from a source node to all its destinations, the delay to each destination will be minimum, and thus the delay of the multicast path (defined as the delay to the farthest destination) will be minimum. The paths can be computed independently, and then merged. In a WDM network, the Shortest Path with Reconfiguration algorithm is also optimum for a single destination. However, if physical multicasting is not possible (i.e., the receivers in the network are not tunable), as routes are computed, links get locked and the topology changes; the tree found is dependent on the order in which the routes to the individual destinations are computed. In this case, the global optimum cannot be decomposed into a number of individual subproblems. If physical multicasting is allowed (i.e., the receivers are tunable), then this algorithm is optimum, because: (i) all transmitters are always free, regardless of the tuning of the receivers, and (ii) if the path from the source node s to destination d_i requires that receiver R_k be tuned to transmitter T_1, and the path from s to destination d_j requires that the same receiver R_k be tuned to transmitter T_2, it is immediate that either $T_1 = T_2$ or the paths from s to T_1 and to T_2 have the same lengths, so that one can again choose $T_1 = T_2$.

4.3.2 Minimum-Cost Routing Heuristic Algorithm

The minimum-cost multicast routing heuristic algorithm presented here is based on the Takahashi-Matsuyama (TM) minimum Steiner tree heuristic [21]. The basic idea in the original TM algorithm is to start building a tree with the source node, and at each iteration add to the tree the destination closest to it. The same idea can be used in the WDM network, but here we use the shortest path with reconfiguration algorithm.

Formally, the minimum-cost heuristic is:

INPUTS: A WDM network, where some transmitters and receivers are locked into a certain topology, and some are free; and a multicast to be routed, characterized by its source node s, and its n destinations $\{d_1, d_2, \ldots, d_n\}$.

OUTPUTS: The updated topology for the WDM network, and the multicast path from the source to the destinations.

ALGORITHM:

Step 1: Add the source node s to the multicast path. Create a virtual node V where the path is "collapsed", i.e., all the nodes in the multicast path are removed from the network and node V "inherits" their transmitters and receivers.

Step 2: Find the shortest path with reconfiguration from V to all the multicast destinations not yet in V. Each path is computed independently from the others, i.e., without taking into account the changes in topology required by the other paths.

Step 3: From all the paths found in step 2, choose the shortest and discard the others. "Collapse" all the nodes in the path into V.

Step 4: If all the destinations have been added to V, stop. The desired multicast path is the result of merging all the paths in V. If there are still destinations not in V, return to step 2.

5 Evaluation of the Multicast Routing Heuristic Algorithms for WDM Networks

In this section, we present an evaluation of the heuristic algorithms proposed in section 4 for multicast routing in a WDM network, using simulation. As done in [22, 23], we first evaluate the performance in a baseline case, chosen arbitrarily, and then determine how variations in the baseline case change the performance measures. We view the WDM network as a backbone network; its nodes can be seen as routers, where the end-user subnetworks are connected.

5.1 The Baseline Case

The baseline case is selected arbitrarily as a "medium-size" network, and is used as a reference point. We first characterize the performance of the baseline case, and then change network parameters and compare the performance to that of the reference point. The following elements characterize the baseline case:

Traffic Model: We assume single-stream sessions, sources and destinations uniformly distributed in the network, stream bandwidths set to 10% of the link bandwidth, number of destinations uniformly distributed between 1 and 10. The number 10% was chosen because we assume the WDM network to be a backbone network; in a well-dimensioned backbone, individual user streams should not occupy a significant fraction of the backbone bandwidth. Sessions arrive according to a Poisson process, and if accepted remain in the network for an exponential period of time. There are no latency constraints.

Network Scenario: 12-node networks, each node having 2 transmitters and 2 receivers. The network starts with a strongly-connected[1] topology generated at random. We consider the following three kinds of networks:

- networks with tunable transmitters (no physical multicast);
- networks with tunable receivers (physical multicast allowed); and
- fixed-topology networks (corresponding to the initial topology of the WDM network, which is a strongly-connected topology generated at random).

Transmitter/receiver costs were set to 0.5, making the cost of a path equal to the number of hops in the path. The distances between the nodes and the WDM star coupler were generated at random, uniformly between 0 and 15; for each node, delays for its transmitters and receivers were set to the distance to the star coupler.

Algorithms Evaluated: For each network scenario, we evaluated the minimum-cost and the minimum-delay heuristic. If the network under consideration allowed physical multicast, it was employed. For the fixed-topology case, the minimum-delay heuristic reduces to the traditional shortest path, and the minimum-cost heuristic to the Takahashi-Matsuyama algorithm.

[1]There is at least one path between any pair of destinations.

Operation of the Network: Each stream is routed as it arrives to the network, considering locked all transmitters and receivers that are currently in use. If the arriving stream cannot be routed, the network considers re-routing as follows: the existing connections plus the newly-arrived request are assumed to be a single session, arriving at an empty network, and the routing algorithm being evaluated is used in sequential order as described in section 4.1. If this session can now be accepted, the existing streams are re-routed and the incoming request is accepted; otherwise, the existing routes are preserved and the arriving stream is blocked. Re-routing is accomplished by momentarily stopping the streams, changing the tuning and the routes, and resuming communication. Audio/video streams can tolerate these short interruptions as long as sufficient data is pre-buffered a the receiver prior to starting the playback; the amount of pre-buffering determines the number of times the stream can be re-routed. If a certain stream cannot be re-routed, this can be easily accommodated by marking the resources used by that stream as locked. In the evaluation presented in this section, we did not assume any limit to the number of times a stream can be re-routed. The reader is referred to [3, 19] for a more complete discussion of the issues in re-routing streams in a WDM network.

Performance Measures: The routing and reconfiguration problem, as formulated in section 3, consists in finding the best route (based on cost or delay) given the incoming session and the network state. However, in a dynamic environment, where sessions come and go, it is interesting to evaluate the blocking probability, i.e., the probability that the session is denied service for lack of resources. Minimizing this probability is a complex problem, and possibly may require taking into account the dynamic behavior of the call arrival process. Cost minimization corresponds really to minimizing the usage of network resources for that stream; if usage of resources is minimized, the blocking probability may be lowered. On the other hand, minimizing delay will reduce the latency experienced by the stream, but may increase the usage of network resources and lead to higher blocking probability. No claims are made that any of these strategies is optimal in relation to blocking probability; in fact, this is an open problem, especially on WDM networks under multicast traffic. With this in mind, we will use the blocking probability as one of the performance measures, together with the average cost/delay for the multicast streams. Another performance measure of interest is the average time between reconfigurations.

The costs and delays of established sessions as a function of the session arrival rate in the baseline case are shown in Figure 2, for the various algorithms and network scenarios. Figure 3 shows the corresponding blocking probability. The main conclusions from Figures 2 and 3 are:

- The average costs and delays for *accepted* streams in networks capable of physical multicasting are the lowest due to their additional flexibility. Both for costs and delays, there is almost a factor of two between the costs and delays of the best and worst algorithms. The relative ranking of the algorithms is what we would intuitively expect.

- Networks where physical multicast is allowed present much higher blocking than networks where it is not allowed. For example, when $\lambda/\mu = 10$, the blocking probability is under 0.2% for the fixed-topology networks and for the networks with no physical multicast, while it is about 30% for networks with physical multicast. For such networks, there is basically no difference between minimum-cost and minimum-delay algorithms as far as the blocking probability is concerned. The main reason for this result is that in the baseline case, the number of transmitters and receivers in the network is the same; therefore, if two receivers are tuned to the same transmitter, there is a transmitter in the network that it not connected and cannot be used. The gain in flexibility does not make up for the lost capacity.

- Minimum-cost algorithms perform better than minimum-delay algorithms, as expected, since they tend to conserve network resources. Also, for the same kind of algorithm (i.e., minimum cost or minimum delay), the WDM networks perform better than the fixed-topology networks. However, it is interesting to note that the blocking probability for the fixed-topology under minimum-cost routing is *lower* than the blocking probability of the WDM network under minimum-delay routing.

The plots in Figure 3 have shown that the blocking probability for the networks where physical multicast is allowed is much higher than when it is not. This seems counter-intuitive: the networks with physical multicast have an *additional* degree of freedom, therefore they should have "better" performance. This "better" performance is seen in the cost and delay of accepted sessions; as discussed before, minimizing the blocking probability is not the objective of minimum cost/delay algorithms. What is happening in the networks where physical multicasting is allowed is that, for each additional receiver tuned to a transmitter, another transmitter in the network will be idle, because the number of transmitters and receivers in the baseline case is the same. Therefore, the number of links in the network is reduced and so is its available aggregate bandwidth. In section 5.4 we will consider networks with more receivers than transmitters.

We have assumed that the traffic has no latency constraints. In general, any routing algorithm can deal with latency constraints by computing the routes, checking the latency, and blocking the request if the latency of the route exceeds the constraint. Now, let us assume that an ideal algorithm capable of minimizing cost subject to a delay constraint were available; we will describe its blocking probability as a function of the latency constraint. Refer to the results of Figure 3. If the latency constraint is very large (i.e., larger than the longest possible path in the network), then our ideal algorithm would reduce to the traditional minimum-cost algorithm and its blocking probability would be the same. As the latency constraint is lowered, the curve will move towards higher blocking probabilities until it reaches the minimum-delay curves. After this point, the algorithm becomes a minimum-delay algorithm, and if the latency constraint is lowered even further, there will be severe blocking.

Figure 4 shows the average time between stream re-routes as a function of the session arrival rate in the baseline case. The time is normalized to the session duration, i.e., an average re-route time of 10 indicates that the average time between re-routing of established streams is 10 times the average session duration (or, on the average, only one session in 10 will suffer rerouting in its lifetime). The figure indicates that only a small fraction of the streams will be re-routed during their lifetimes, for reasonable traffic loads. Moreover, it is very unlikely that a stream will will be rerouted more than once. Re-routing is more severe in networks where physical multicast is allowed, due to the high blocking. We also determined that the average path change when a stream is re-routed is very small; the average change in cost is under 0.2 hops for the topologies with no physical multicast, and under 0.6 hops for the topologies with physical multicast.

5.2 Changing the Network Size

In this section, we investigate the effects of increasing the network size. The WDM network can be seen as a backbone network, responsible for serving a certain user population. If the capacity of this network is to be increased, there are two options: (i) increase the number of nodes, thus reducing the fraction of the user population that a node needs to support, and/or (ii) increase the number of transmitters and receivers per node, thus increasing the node's available bandwidth. To investigate these alternatives, we consider the following two variations to the baseline case:

- Increase the number of transmitters and receivers per node from 2 to 3, keeping the number of nodes fixed at 12.
- Increase the number of nodes from 12 to 18, keeping the number of transmitters and receivers per node fixed at 2.

Note that, in both cases, we increase the total number of transmitters and receivers in the network from 24 to 36. The traffic model is kept the same as the baseline case.

Figure 5 shows the blocking probability as a function of the session arrival rate for a 12-node network, with 3 transmitters and 3 receivers per node. We note that the difference in performance between the fixed-topology network and the WDM network has decreased considerably. This is explained by the fact that the increased node degree leads to a richer topology, with shorter paths; the additional degree of freedom of rearranging links in the WDM network is less important in this scenario. We also observe that re-routing happens less frequently than in the baseline case.

Figure 6 shows the blocking probability as a function of the session arrival rate for a 18-node network, with 2 transmitters and receivers per node, under the same traffic as the baseline case. The same qualitative comments about the algorithms made in the baseline case can be repeated here. We note that the advantage of the WDM network over the fixed-topology network has increased. Another observation is that since the paths get "longer", the average change in cost due to re-routing can be as high as 1 hop for the networks where physical multicast is allowed, and 0.6 hops where it is not.

Figure 7 shows a comparison of the blocking probability for the two cases, and shows clearly that it is much better to increase the network capacity by increasing the number of transmitters/receivers per node, rather than increasing the number of nodes; this is a well-known result for fixed-topology graphs. The reason is that the increased number of links decreases the average path length (besides, of course, increasing the capacity), thus reducing the load in the network. On the other hand, increasing the number of nodes *increases* the path length, thus offsetting some of the capacity increase. Similar results are found for minimum-delay algorithms.

5.3 Keeping the Network Connected

In the simulation runs for the baseline case, we did not require the network to stay connected as its topology changes. In this section, we impose the additional constraint that at every invocation of the Shortest Path with Reconfiguration Algorithm, the resulting network stays strongly-connected.

The idea is as follows: we start with the optical network in a strongly-connected topology. Note that all transmitters and receivers are free, but instead of assuming that they are unallocated, we assume that they are tuned to form a strongly-connected topology. If, during operation, the Shortest-Path with Reconfiguration Algorithm recommends a topology change, the new topology is tested for connectivity, using the Floyd-Warshall algorithm [20] to compute its minimum distance matrix[2]; if the matrix contains any infinite entry, the topology is not connected. If the new topology is strongly-connected, the reconfiguration is accepted. Otherwise, we search for a *secondary reconfiguration* that will make the network connected again, by attempting to connect nodes with an infinite entry in the distance matrix. If a secondary reconfiguration is found, the topology is changed; otherwise, the original reconfiguration proposed by the Shortest-Path Algorithm is rejected and the stream routed through the current topology, if possible. In [3] we describe in detail this algorithm.

Figure 8 shows a comparison of the blocking probability between networks that stay connected and networks that are allowed to have arbitrary topologies, for minimum cost algorithms. The figure indicates that while there is basically no change in the results for networks where physical multicast is not possible, there is a significant improvement for networks where it is used, because less transmitters are left unassigned. Of course, the average costs and delays for the accepted sessions increase. Another observation is that for networks allowing physical multicast, keeping the topology connected increases the average time between re-routes and decreases the change in cost, especially for minimum-delay algorithms.

5.4 Changing the Number of Receivers

In the previous sections, we have shown that allowing physical multicast in a WDM network where number of transmitters is equal to the number of receivers greatly degrades the blocking probability. Keeping the network connected offsets some of this degradation, but it is still severe. However, we expect that in a network where receivers are abundant (i.e., more receivers than transmitters), allowing physical multicast should improve the performance.

To investigate this fact, we obtained the blocking probability for a number of 12-node networks, keeping the number of transmitters per node fixed at 2, and varying the number of receivers per node from 2 to 8. The session arrival rate was kept fixed, and no connectivity

[2]Element (i, j) of the minimum distance matrix is the length of the shortest path between nodes i and j.

constraints were imposed on the network. The blocking probability results (as a function of the number of receivers) are shown in Figure 9. The baseline case corresponds to 2 receivers per node.

Figure 9 shows that, as expected, networks where physical multicast is not possible cannot make effective use of extra receivers. When physical multicast is allowed, adding receivers will improve the performance. For this particular case, the point where the performance of the two networks becomes the same is between 5 and 6 receivers. For higher loads, the point of equal performance shifts to a lower number (for example, for $\lambda/\mu = 40$, the crossover point is between 4 and 5 receivers).

We also investigated the role played by the transmitter/receiver costs in the blocking probability, as the number of receivers change. In the baseline case, transmitter and receiver costs are the same (equal to 0.5). We found that, for networks where physical multicasting is not allowed, the blocking probability not sensitive to the allocation of costs. However, in networks where physical multicasting is allowed, making the receivers "expensive" (and thus making the physical multicast "expensive"), improves the blocking probability. Detailed results can be found in [3].

6 Discussion and Conclusions

In this paper, we considered the problem of routing multimedia streams in a WDM network, taking into account their requirements of bandwidth, multipoint communications, and latency, and making use of the additional degree of freedom (the topology) of the network. We proposed and evaluated heuristic algorithms to solve this problem.

The main conclusion from the evaluation of multicast routing algorithms in the WDM network concerns the use of physical multicasting in a network where receivers are tunable. The unrestricted use of physical multicasting can decrease the cost and delay of accepted sessions, but at the price of a much higher blocking probability. The previous sections have shown that, by keeping the network connected and/or by properly assigning link costs, the penalty paid in performance can be reduced. However, the blocking probability is still much higher than in the case where no physical multicasting is allowed, except when the number of receivers is large. The fundamental reason for this behavior is simply that, by allowing physical multicasting, the network can be re-arranged in a topology where all receivers are in use, but some transmitters are not connected to any receiver. This represents a net *decrease* in network capacity, and leads to increased blocking probability. Keeping the network strongly connected and/or shifting the cost to the receivers improves the performance because less transmitters are allowed to remain disconnected. By the same token, when there are plenty of receivers, it is unlikely that transmitters will be disconnected. In summary:

- Networks where physical multicasting is allowed are able to make use of extra receivers as they are added; however, performance with a small number of receivers per node is very low; techniques such as keeping the network strongly connected and/or properly assigning transmitter/receiver costs can help, but not much.
- Networks where physical multicasting is not allowed have much better performance when the number of receivers is small, but cannot make use of additional receivers.

These observations indicate that the best way to use physical multicasting is:

1. The network should be kept strongly connected at all times.
2. At all times, there should be *at least one* receiver connected to each transmitter.
3. Physical multicasting can be used as long as it does not violate rules 1 and 2 above. Note that this is true regardless of the type of traffic (unicast or multicast) in the network. Even unicast traffic can benefit from the extra flexibility of allowing physical multicasting, as long as rules 1 and 2 are satisfied.

By using the above rules, we can guarantee good performance both when the number of receivers is small and when it is large. Note that, if the number of receivers is equal to the number of transmitters, no physical multicasting will be allowed.

Based on the results of this paper, we derive the following set of guidelines for designers/implementors of WDM networks:

- If the network has tunable transmitters and fixed receivers, the total number of transmitters and receivers in the network should be the same. To upgrade a node's capacity, one has to add transmitters and receivers to it in pairs.

- It is recommended that the network be built with tunable receivers and fixed transmitters. In this case, a node can be upgraded just by adding receivers. When operating the network, physical multicasting should be allowed only if there are more receivers than transmitters, as described above.

- Under multicast traffic, the minimum cost heuristics achieve lower blocking. It is an open problem to determine the role, if any, of the latency constraints.

References

1. L. Kazovsky et al, WDM Local Area Networks, *IEEE LTS*, Vol. 3, No. 2, pp. 8-15, May 1992.

2. W. B. Jones, *Introduction to Optical Fiber Communication Systems*, Holt, Rinehart and Winston, New York, 1988.

3. C. Noronha and F. Tobagi, Routing of Streams in WDM Reconfigurable Networks, Technical Report CSL-TR-94-649, Computer Systems Laboratory, Stanford University, December 1994.

4. Y. Birk et al, Bus-Oriented Interconnection Topologies for Single-Hop Communication among Multi-Transceiver Stations, *Proc. IEEE INFOCOM '88*, New Orleans, LA, pp. 558-567, March 1988.

5. M. Mehdi Nassehi et al, Fiber Optic Configurations for Local Area Networks, *IEEE Journal on Selected Areas in Communications*, Vol. 3, No. 6, pp. 941-949, November 1985.

6. K. Kobayashi and I. Mito, Single Frequency and Tunable Laser Diodes, *Journal of Lightwave Tech.*, Vol. LT-6, No. 11, pp. 1623-1633, November 1988.

7. T. P. Lee and C. E. Zah, Wavelength-Tunable and Single-Frequency Semiconductor Lasers for Photonic Communications Networks, *IEEE Communications Mag.*, pp. 42-51, October 1989.

8. D.A. Smith et al, Integrated-Optic Acoustically-Tunable Filters for WDM Networks, *IEEE Journal on Sel. Areas in Comm.*, Vol. 8, No. 6, pp. 1151-1159, August 1990.

9. H. Kobrinski and K.-W. Cheung, Wavelength-Tunable Optical Filters: Applications and Technologies, *IEEE Communications Mag.*, pp. 53-63, October 1989.

10. B. Mukherjee, WDM-based local lightwave networks. I. Single-hop systems, *IEEE NETWORK*, Vol. 6, No. 3, pp. 12-27, May 1992.

11. B. Mukherjee, WDM-based local lightwave networks. II. Multihop systems, *IEEE NETWORK*, Vol. 6, No. 4, pp. 20-32, July 1992.

12. J. Bannister, L. Fratta, M. Gerla, Topological Design of the Wavelength-Division Optical Network, *Proc. IEEE INFOCOM 1990*, San Francisco, June 3-7, 1990.

13. J. Bannister, L. Fratta, M. Gerla, Routing in Large Metropolitan Area Networks Based on Wavelength-Division Multiplexing Technology, in *High-Capacity Local and Metropolitan Area Networks - Architecture and Performance Issues*, edited by G. Pujolle, Springer-Verlag, 1990.

14. J.-F. Labourdette and A. Acampora, Logically Rearrangeable Multihop Lightwave Networks, *IEEE Trans. on Comm.*, Vol. 39, No. 8, pp. 1223-30, August 1991.

15. J.-F. Labourdette and A. Acampora, Partially Reconfigurable Multihop Lightwave Networks, *Proc. IEEE GLOBECOM 1990*, San Diego, CA, December 2-5, 1990.

16. J. Auerbach and R. Pankaj, Use of Delegated Tuning and Forwarding in Wavelength Division Multiple Access Networks, *IEEE Trans. on Comm.*, Vol. 43, No. 1, January 1995.

17. C. Noronha and F. Tobagi, Optimum Routing of Multicast Streams, *Proc. IEEE INFOCOM '94*, Toronto, Canada, pp. 865-873, June 1994.

18. C. Noronha and F. Tobagi, Optimum Routing of Multicast Audio and Video Streams in Communications Networks, Technical Report CSL-TR-94-618, Computer Systems Laboratory, Stanford University, April 1994.

19. C. Noronha and F. Tobagi, Routing of Multimedia Streams in Reconfigurable WDM Optical Networks, *Journal of High-Speed Networks*, Vol. 4, No. 2, pp. 133-153, 1995.

20. D. Bertsekas, and R. Gallager, *Data Networks*, Prentice-Hall, New Jersey, 1987.

21. H. Takahashi and A. Matsuyama, An Approximate Solution for the Steiner Problem in Graphs, *Math. Japonica* 6, pp. 573-577, 1980.

22. C. Noronha and F. Tobagi, Evaluation of Multicast Routing Algorithms for Multimedia Streams, *IEEE International Telecommunications Symposium*, Rio de Janeiro, Brazil, August 1994.

23. C. Noronha and F. Tobagi, Evaluation of Multicast Routing Algorithms for Multimedia Streams, Technical Report CSL-TR-94-619, Computer Systems Laboratory, Stanford University, April 1994.

Biographies

Ciro A. Noronha Jr. received the Electronics Engineer and Master of Electronics Engineering degrees from the Instituto Tecnológico de Aeronáutica, São José dos Campos, Brazil, in 1985 and 1988, respectively, and the Ph.D. degree in Electrical Engineering from Stanford University in 1995.

From 1986 to 1987 he was a Junior Design Engineer at Elebra Telecon, São Paulo, working with data communications. From 1987 to 1989 he was an Auxiliary Professor with the Instituto Tecnológico de Aeronáutica, São José dos Campos, and from 1990 to 1994 he was a Research Assistant at Stanford University. He currently a Member of the Technical Staff at Optivision, Inc., where he is responsible for the design and implementation of the routing protocols for Optivision's Optical Crossbar Switch. He is also responsible for Optivision's MPEG-1 and MPEG-2 live video transmission and conferencing systems.

Fouad A. Tobagi received the Engineering Degree from Ecole Centrale des Arts et Manufactures, Paris, France, in 1970 and the M.S. and Ph.D. degrees in computer science from the University of California, Los Angeles, in 1971 and 1974, respectively.

From 1974 to 1978, he was a Research Staff Project Manager with the DARPA project at the Computer Science Department, University of California, Los Angeles, and engaged in modeling, analysis, and measurements of packet radio systems. In June 1978, he joined the faculty of the School of Engineering at Stanford University, Stanford, California, where he is currently Professor of Electrical Engineering and by courtesy Computer Science. He is also a Co-Founder of Starlight Networks, a new venture concerned with multimedia networking, where he is serving as the Chief Technical Officer. His current research interests include fast packet switching, broadband integrated services digital networks, ATM, multimedia communications, video servers, modeling and performance evaluation, and VLSI implementation of network components.

Dr. Tobagi is a fellow of the IEEE for his contributions to the field of Computer Communications and Local Area Networks. He is the winner of the 1981 Leonard G. Abraham Prize Paper Award in the field of Communications Systems and co-winner of the IEEE Communications Society 1984 Magazine Prize Paper Award. He has served as Editor for several technical journals and is is currently serving on the editorial boards of the Journal of High Speed Networks, the ACM Journal of Multimedia Systems, the Journal of Multimedia Tools and Applications, and the Journal of Wireless Networks.

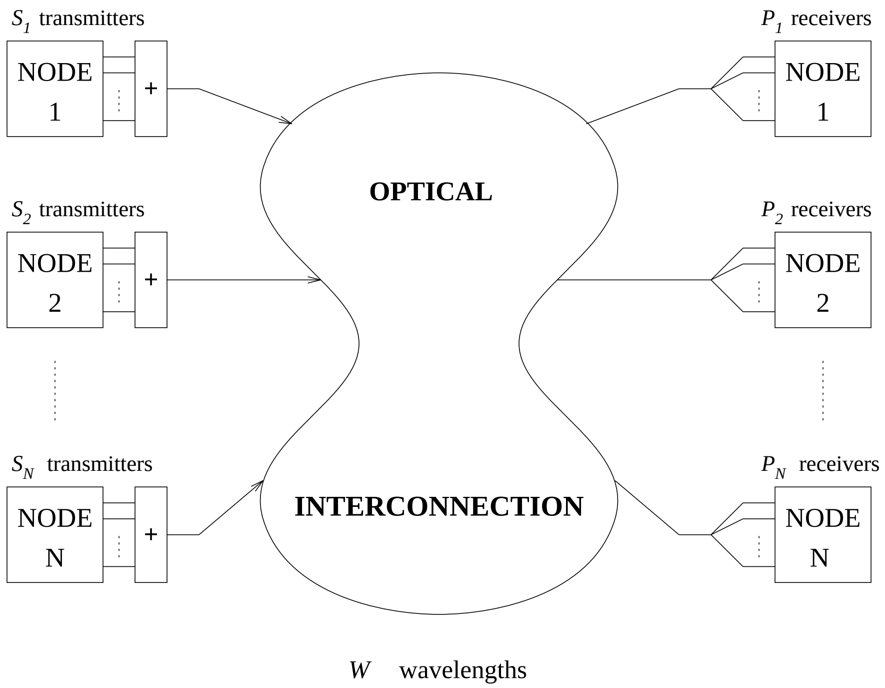

Figure 1: The WDM Network

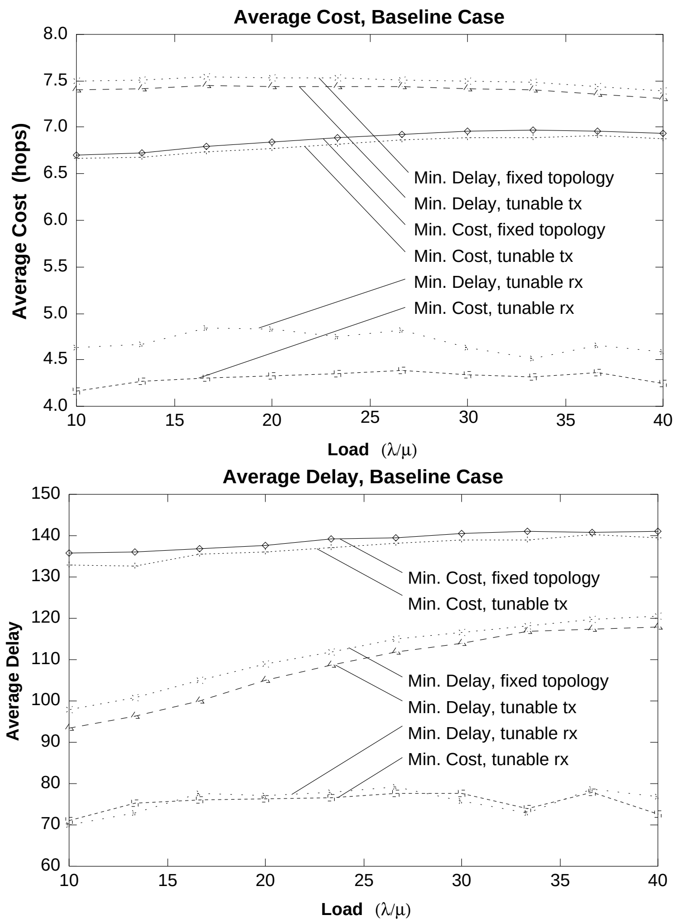

Figure 2: Comparison of costs and delays for established sessions in the baseline case: single-multicast sessions, stream bandwidth 10% of the link capacity, 50,000 routes/point

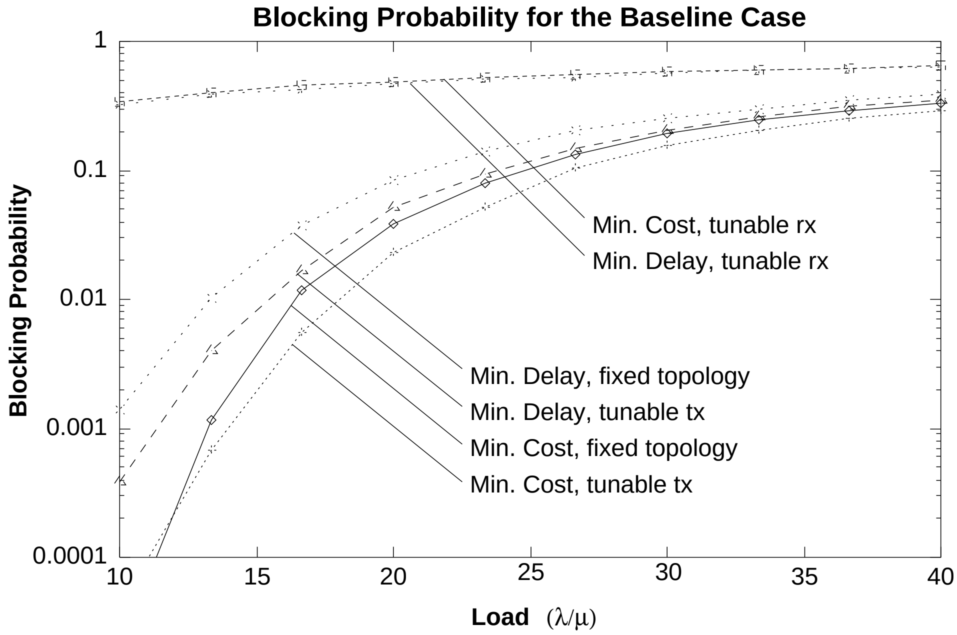

Figure 3: Blocking probability for the baseline case: single-multicast sessions, stream bandwidth 10% of the link capacity, 50,000 routes/point

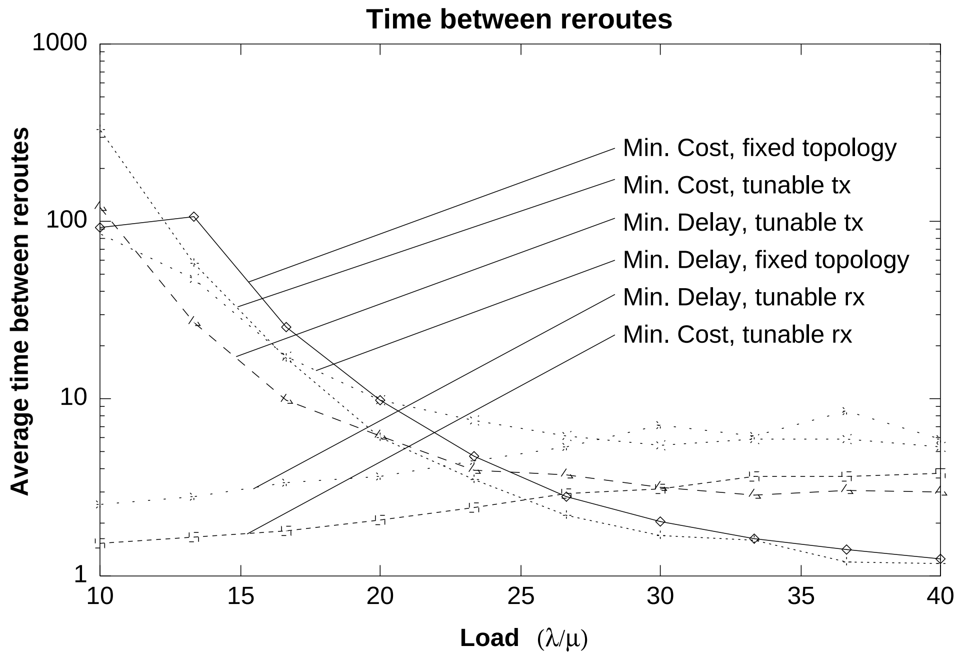

Figure 4: Average time between stream re-routes, baseline case

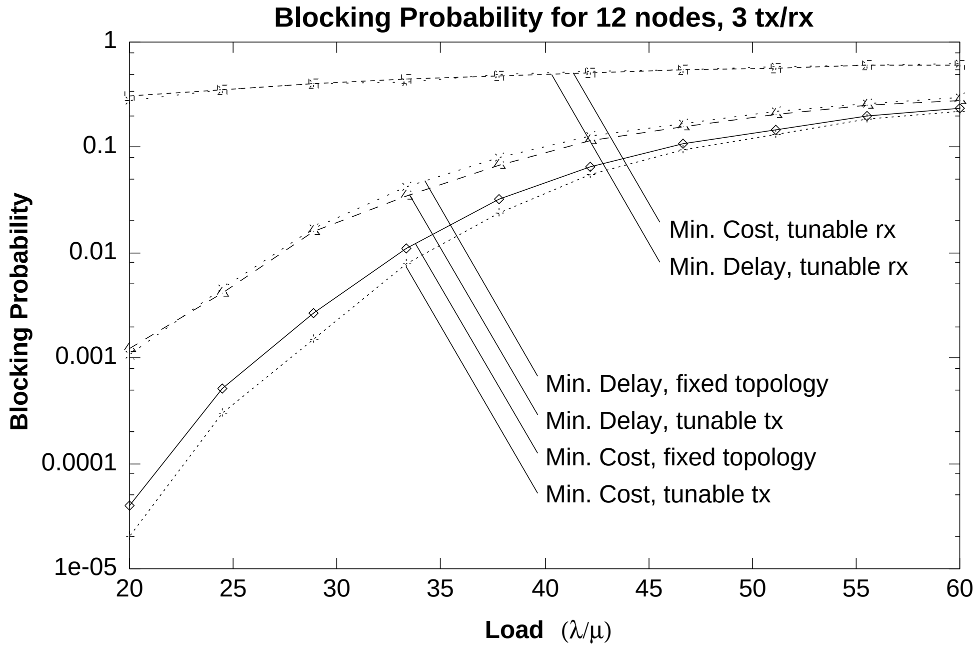

Figure 5: Blocking probability for 12-node networks, 3 transmitters/receivers per node, 50,000 routes/point

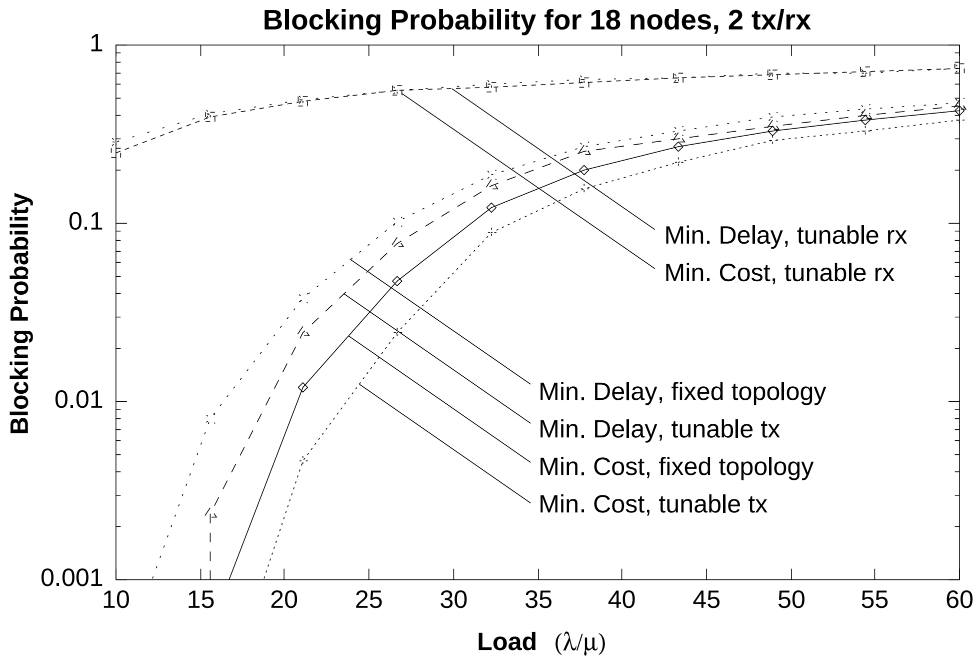

Figure 6: Blocking probability for 18-node networks, 2 transmitters/receivers per node, 15,000 routes/point

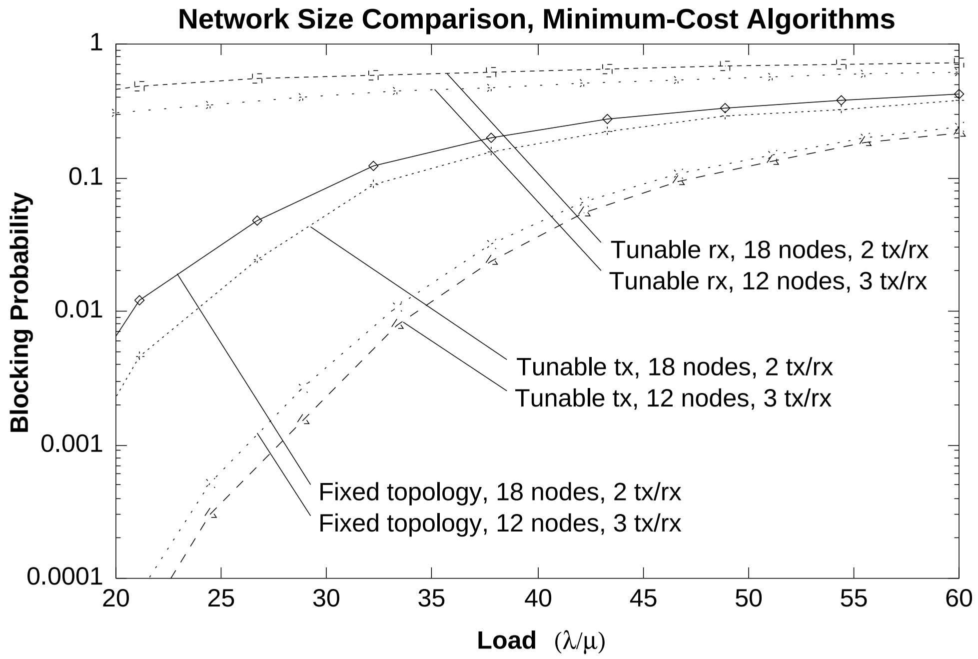

Figure 7: Comparison of the blocking probability for minimum-cost algorithms, when the size of the network is increased

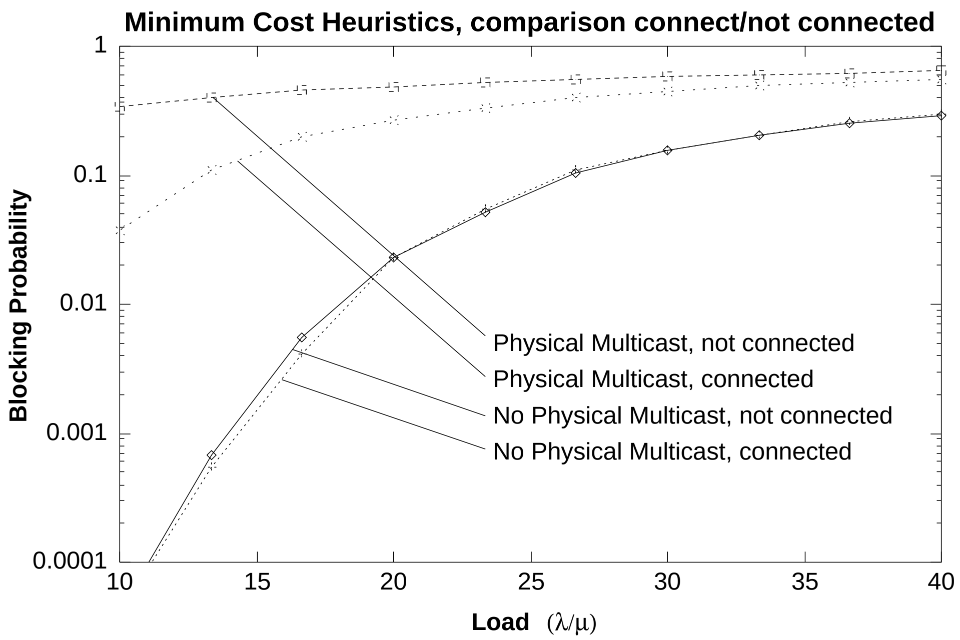

Figure 8: Comparison of blocking probability for minimum cost heuristics, when the network is kept connected versus when it is not

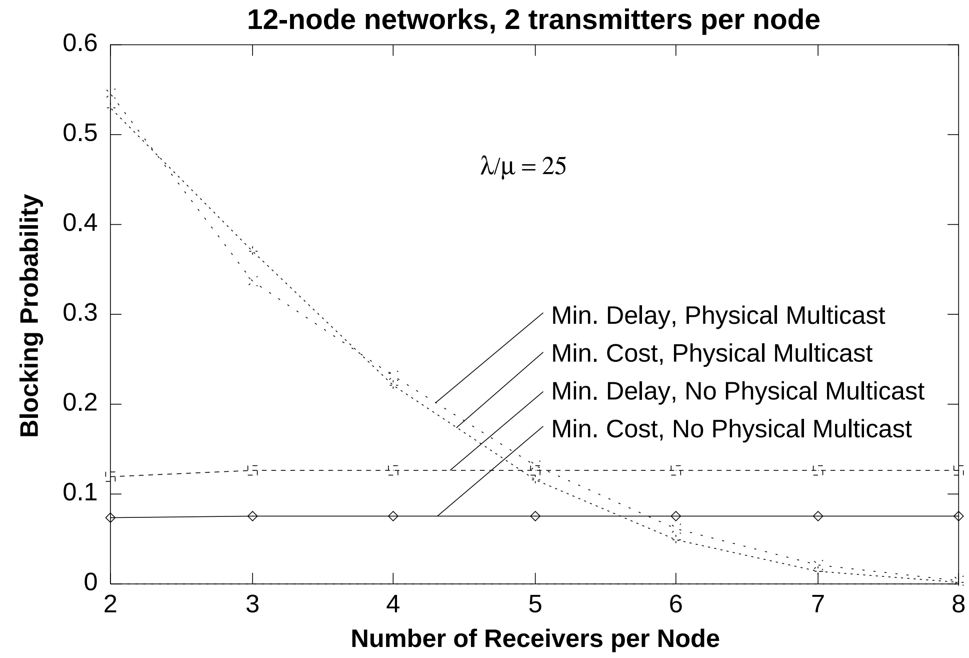

Figure 9: Increasing the number of receivers per node, keeping the number of transmitters fixed, 20,000 routes/point