Routing of Multimedia 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@bodega.stanford.edu
Phone: +1 (415) 723-0803
Fax: + 1 (415) 725-7398

Abstract

Multimedia streams have different requirements than those of traditional data traffic, especially in terms of bandwidth; these requirements have to be taken into account when routing such streams. WDM local-area networks with tunable transceivers have an additional degree of freedom over their electronic counterparts: their topology can be dynamically changed. In this paper, we study the problem of routing multimedia streams in WDM networks, taking into account their traffic requirements and making use of the network's ability to dynamically change the topology. We show that the reconfiguration and routing problem can be written as a linear integer programming problem. Since the exact solution to this problem is complex, we present a simpler heuristic, and show that it provides good results, thus obviating

*This work was in part supported by NASA under grant NAGW-419, by NSF 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.

the need for the more complicated exact solution. Finally, we evaluate the performance of the WDM network in a realistic traffic scenario, and show that it compares favorably with that of a centralized switch.
Keywords: WDM networks, multimedia streams, routing algorithms, performance evaluation.

1 Introduction

Multimedia represents the integration of a variety of media, such as data, video, audio and still images. The traffic underlying networked multimedia applications has different requirements than that underlying traditional data applications, and networks carrying such traffic should take into account the specific requirements of multimedia streams.

The main difference between the traffic generated by traditional data applications and the traffic from multimedia streams is one of bandwidth: multimedia streams use relatively high bandwidth in a continuous basis for long periods of time, while the average bandwidth used by data applications is low. 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. Therefore, when routing streams, the network must keep track of the bandwidth usage of each link.

Multimedia applications generate traffic in *sessions*; a session is a group of streams which are logically related. A stream is a continuous flow of information from a source node to one or more destinations. The network's *routing algorithm* is then responsible for finding routes for each of the streams in the session with enough free bandwidth to support them. When the session (or any of its components) terminates, the network de-allocates the corresponding bandwidth. Although multimedia applications are expected to make heavy use of multicasting, for this paper we consider only unicast traffic (one destination only), leaving multicast for a future paper.

Due to its low attenuation (less than 0.2 dB/km) and very high bandwidth, fiber has became the medium of choice for point-to-point links at high speeds, for any distance over ≈ 100 m. Using Wavelength-Division Multiplexing (WDM), many channels can be created in the fiber. A tunable transmitter allows the network node to select one of these channels for sending data, and to dynamically change this selection. The number of wavelengths addressable depends on the technology used to implement the tunable transmitter. Similar comments can be made for tunable receivers. By making the combined light signal from the transmitters available to all receivers (by using, for example, a WDM star coupler), the interconnection pattern between nodes will be defined by the tuning of transmitters and receivers to specific wavelengths, and can be dynamically changed. In a traditional (fixed-topology) network, given a multimedia session, the routing algorithm is responsible for finding routes for each of its components, satisfying the session 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.

In this paper, we consider the problem of dynamically routing multimedia streams in a WDM local or metropolitan-area network, taking into account their bandwidth requirements. In section 2, we give a formal problem definition, and indicate that the optimum routing and reconfiguration problem can be written as an integer linear programming problem. In section 3 we briefly describe the previous work in the area. Since the optimum solution is complex and, in any case, the problem is NP-complete, in section 4 we give an heuristic algorithm which will find a sub-optimal solution. This algorithm is evaluated in section 5, where we derive an upper bound on the performance of any algorithm, and show that the heuristic produces results close to this upper bound, thus obviating the need to pursue the optimum solution. Finally, in section 6, we evaluate the performance of the WDM reconfigurable network in a realistic environment, and show that its performance is comparable to that of a centralized ATM switch. Our conclusions are presented in section 7.

2 Problem Definition

In this section, we present a formal definition for the following problem: "Given a WDM optical network and a multimedia session, find the topology (i.e., the tuning of the transmitters and/or receivers) and routes for the components of this session that satisfy its bandwidth requirements and optimize a certain objective function". We first describe the environment (network and traffic models), and then present the problem formulation.

2.1 The Environment

2.1.1 The Network Model

The WDM network is a set of nodes, each equipped with a number of optical transmitters and receivers, and an optical interconnection [1]:

Optical Transmitters: Each transmitter sends data using a specific wavelength. If the transmitter is tunable, it can use different wavelengths at different times. If two different transmitters send data in the same wavelength, their signals are mixed and the data cannot be received. An important parameter is the output optical power of the transmitter; it determines how many receivers can be reached by that transmitter.

Optical Interconnection: Responsible for combining the light from the transmitters and splitting the resulting signal between the receivers. Note that the optical interconnection does not necessarily deliver the signal from all transmitters to all receivers - a given receiver might have access to only a subset of the transmitters. The most common optical interconnection, for the local and metropolitan area network environments, is the WDM star, which equally divides the light from each of the transmitters to each of the receivers. Other topologies for the optical interconnection, such as trees, are possible; however, the splitting of the transmitted power between the receivers will be unequal.

Optical Receivers: The optical receiver is able to extract one specific wavelength from the combined signal coming from the optical transmitters, and receive the data from it. An optical receiver can be either fixed or tunable. An important parameter is the *receiver sensitivity*, defined as the minimum optical power at the input of the receiver required to guarantee a bit error ratio of 10^{-9} or better. Given the transmitter power and the distances, the sensitivity will determine how many receivers can be reached by that transmitter. Many receivers can be simultaneously tuned to a single wavelength, providing physical multicasting.

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. We assume that the total number of optical transmitters and receivers in the network is the same and will be denoted by K (i.e., $\sum_{i=1}^{N} S_i = \sum_{i=1}^{N} P_i = K$). This is reasonable under unicast traffic because there is an one-to-one correspondence between transmitters and 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.

- The number of distinct wavelengths, denoted by W, is larger than the number of transmitters/receivers in the network ($W \geq K$). Given the bandwidth of the fiber, this is a reasonable assumption for a small to moderate size network (i.e., up to 100 nodes).

- At any time, only one transmitter and only one receiver can be tuned to a given wavelength. We do not consider multiple-access operation (i.e., many transmitters tuned to the same wavelength) or physical multicast (i.e., many receivers tuned to the same wavelength). The network operates in a point-to-point, store-and-forward fashion.

- Any given transmitter can be connected to any given receiver - there are no tunability restrictions. For the purposes of this paper, it does not matter which component (the transmitters or the receivers) is tunable.

- Usually, S_i and P_i are much less than N. Therefore, each node will have direct connectivity to a (typically small) subset of nodes.

2.1.2 The Traffic Model

For the traffic, we assume that user requests come in *sessions*. A session is a group of streams that are logically related. We denote by T the number of streams in the session; for the evaluation, we assume that T is a fixed number. Stream i, $i = 1, \ldots, T$, is characterized by its source s_i, its destination d_i, and its bandwidth requirement r_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.

2.2 The Problem Formulation

The problem under consideration can be stated as: "Given the network state (i.e., the routes for the sessions already established) and a new session with T streams, each stream being characterized by its source, destination and bandwidth requirement, find the new 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 network topology, which defines which transmitters are connected (tuned) to which receivers, and (ii) the assignment of routes given the topology. If multiple solutions exist for a reconfiguration/routing problem, the objective function is the criterion used to select the "best" one. The following are possible objective functions:

- minimize the average number of hops in the path.

- minimize the maximum flow over all links.

Both measures have been shown to be roughly equivalent [2, page 356]. Since we focus in the LAN/WAN environment, for the remainder of this paper we will use the average hop count as our optimization criterion.

2.3 Solving the Reconfiguration and Routing Problem

For a given network topology, the problem of routing multimedia streams satisfying their bandwidth requirements and minimizing the average number of hops in the path corresponds to the well-known multicommodity flow problem, with a linear objective function, which can be exactly solved by linear integer programming [3]. However, in a WDM network, the topology itself is a variable, and the flow conservation equations become non-linear, because they involve the product of two sets of free variables (the flows and the topology).

In the Appendix, we show that, by adding new free variables to represent this product between the flow and the topology, and additional constraints, the problem of optimally routing a multimedia session on a WDM network can be formulated as a *linear* integer programming problem, which can be solved by standard techniques such as the branch-and-bound method [4].

3 Previous Work

A comprehensive review of the work in the field of WDM networks can be found in [5] and [6]. Here we present a short summary; as done in [5, 6], we classify the existing reconfiguration and routing algorithms as *single-hop*, where the source and the destination share a wavelength for the duration of the data transfer, and *multi-hop*, where the data might potentially go from node to node until it reaches the destination.

3.1 Single-Hop Algorithms

Habbab et al [7] and later Mehravari [8] considered a network where the number of distinct wavelengths is much smaller than the number of stations. Each station has one tunable receiver and one tunable transmitter, and both are capable of addressing all the wavelengths in the network. Coordination between transmitters and receivers is achieved by reserving one wavelength for control; all idle nodes keep their receivers tuned to this wavelength. Multiple-access schemes are used both in the control and in the data channels.

Chlamtac and Ganz [9] and later Ganz and Koren [10] considered a scenario where all the stations are synchronized, the transmitters are fixed and the receivers are tunable. All packets arrive aligned at the star coupler. Coordination between transmitters and receivers is achieved by having a common "tuning schedule", known by all nodes. Control algorithms and approximate analysis based on Markov chains are presented.

In summary, the work done in single-hop algorithms assumes that the tuning of transmitters and/or receivers can be very fast, and that the network either uses a multiple-access scheme (which is difficult to implement efficiently in optics) or is synchronized (which might be difficult to achieve at high speeds).

3.2 Multi-Hop Algorithms

In [11, 12, 13, 14] 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 light signal from all transmitters is available to all receivers. The differences are in the following areas:

(i) Objective function to be optimized: in [11], 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 [12], 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 [13, 14], the objective function is to minimize the maximum flow over all links.

(ii) Additional constraints in the optimization: in [14] 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 [11, 12] the objective function is non-linear, and the authors resort to the "simulated annealing" method to search for a sub-optimal solution. In [13, 14] the authors present an 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 work in the area of multi-hop WDM networks has focused in long-term topological design of the topology, assuming Poisson traffic. In this paper, we consider that the network reconfigures on a connection-by-connection basis, allowing it to respond to instantaneous traffic variations, but without the need for a multiple-access scheme or network-wide synchronization. Moreover, we consider the routing of multimedia streams; since they require reserved bandwidth in the links (and are much less bursty than the traffic generated by data applications), the delay is independent of the traffic in the links, allowing the use of a linear objective function (the hop count).

4 A Heuristic Solution to the Reconfiguration and Routing Problem

The reconfiguration and routing problem, as formulated in section 2, is NP-complete, and the exact optimum solution has (in the worst case) exponential run time. In this section, we present a simpler heuristic solution to the problem.

Given a session with one or more streams, we seek to find the network topology and the routes for this session. From a high level point of view, the heuristic solution proposed here starts with an arbitrary initial topology (i.e., all the transmitters and receivers are connected in a certain fashion), 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 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 heuristic.

4.1 The Shortest Path with Reconfiguration Algorithm

Given a source node, Dijkstra's algorithm builds a shortest path *tree* from that node. The tree starts with the source node, and at each iteration a node is added to it; the paths in the tree are the shortest from the source. When used to find the shortest path between a particular pair of nodes, the algorithm terminates when the destination node is added to the tree, at which point only the path between the source and the destination nodes is retained, and the remainder of the tree is discarded.

We seek 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 - transmitters and receivers might be already in use. In general, we classify the transmitters and receivers in the network either as *free* or *locked*, and 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, as long as there is enough bandwidth available. The algorithm described below does exactly this: given the WDM network in a certain 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. In the Appendix, we give a formal description of the algorithm.

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.

NOTE: At most one reconfiguration is needed to obtain the shortest path (and the algorithm above finds it). This can easily be shown by contradiction: assume that the shortest path between nodes S and R requires that node A be reconfigured to connect to node B, and node C be reconfigured to connect to node D. In the absence of tuning constraints, we can reconfigure node A to connect directly to node D, finding a path that is shorter, which contradicts the initial hypothesis.

4.2 The Reconfiguration and Routing Heuristic

Given a session, the basic idea behind the reconfiguration and routing heuristic is to take an arbitrary initial topology, and apply the shortest path with reconfiguration to each of the components of the session, in decreasing order of bandwidth. The steps are:

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 and r_j is the bandwidth requirement for stream j. 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 above for this stream. If successful, update the 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. As a part of this process, the routes are also found, if all invocations of the shortest path with reconfiguration algorithm in step 3 are successful; otherwise, the heuristic fails and declares the problem infeasible.

Note that, for a given network topology, routing a session using this topology becomes the traditional multicommodity flow problem. The branch-and-bound (see, for example, [4]) method can be used to solve this linear integer programming problem. Specific features of the problem can be used to prune the search space and speed-up the solution, as proposed by Crowder et al [15]. In [16] we present a set of enhanced value-fixing rules to accomplish this pruning. Therefore, one can further optimize the solution by using the topology found by the heuristic, disregarding the routes found in step 3, and re-route the session as described. In some cases, by doing this it is possible to solve a problem declared infeasible by the heuristic. For all the results presented in this paper, we have re-optimized the routes using integer programming.

5 Evaluation of The Reconfiguration and Routing Heuristic

In this section, we present an evaluation of the reconfiguration and routing heuristic described in the previous section. Ideally, one would compare the results of the heuristic with the exact (optimum) solution; however, we derive an upper bound in performance, which is much simpler to compute than the optimum, and use this bound in the evaluation. We also compare the performance of the WDM reconfigurable network with a fixed-topology network with the same number of nodes and links; for the evaluation, we chose the ShuffleNet [17]. The routes in the ShuffleNet were computed using linear integer programming [3, 4].

5.1 Evaluation Scenarios and Performance Measures

The first step in the evaluation is defining the evaluation scenarios and performance measures under which the algorithms are to be compared:

Evaluation Scenarios

For the evaluation, we consider networks with $N = 8$ nodes; each node has 2 optical transmitters and 2 optical receivers ($K = 16$). We consider the routing of a single session on an idle network (this is equivalent to making the session arrival rate, λ, much lower than the average session duration, $1/\mu$). The session is composed of T streams; T is fixed at 10, 15 and 20. The sources and destinations of the streams are uniformly distributed over the network. The bandwidth requirement for each stream is generated at random between 0 and 100% of the link bandwidth, using the following bimodal distribution (m is the average bandwidth requirement, expressed as a fraction of the link bandwidth V):

$$p_B(b) = \begin{cases} \frac{1-m}{m} & \text{if } b < m \\ \frac{m}{1-m} & \text{if } b \geq m \end{cases} \tag{1}$$

The average bandwidth required by the session, as a fraction of the total bandwidth in the network, is given by mT/K; we denote this quantity as the *Offered Load* to the network.

Performance Measures

The most basic performance measure is the *Session Acceptance Probability*, which is the probability that there are enough resources in the network to accept a session and route it, i.e., the feasible region of the optimization problem described in section 2 is not empty. Since we seek to minimize the average path length (in hops), this is another useful performance measure. Note that these two performance measures are related: for a given algorithm, the average path length indicates the usage of network resources when routing a session. If this value is high, it is likely that the acceptance probability will be low.

5.2 An Upper Bound on the Session Acceptance Probability

Given a session composed of T streams as described above, it might be impossible to route this session, regardless of the network reconfiguration algorithm. In this section, we establish a necessary condition for a session to be accepted. If this condition is false, the session cannot be accepted; for any evaluation scenario, the fraction of the sessions meeting this condition constitutes an upper bound on the session acceptance probability.

If a session is to be accepted and routed, for each node in the network there should be at least one way of distributing the streams that originate from it among its transmitters. Similarly, there should be at least one way of distributing the streams that are destined to it among its receivers. For example, it is not possible to have three streams requesting 60%

of a link's bandwidth originating at a node that has only two transmitters, although the three streams combined request less bandwidth than the total available, because a stream cannot be split. Formally, a **necessary** condition for the existence of a solution for the routing/reconfiguration problem is that, for every node k, $k = 1, \ldots, N$, at least one feasible solution is found for each of the problems below:

Problem 1: define $\mathcal{A}_k$ to be the set of streams that originate at node k; find a set α_{ij}, $i \in \mathcal{A}_k$, $j = 1 \ldots, S_k$ such that:

$$\sum_{i \in \mathcal{A}_k} \alpha_{ij} b_i \leq V, \quad j = 1, \ldots, S_k \tag{2}$$

$$\sum_{j=1}^{S_k} \alpha_{ij} = 1, \quad \alpha_{ij} \text{ is binary}, \quad \forall i \in \mathcal{A}_k$$

Problem 2: define $\mathcal{B}_k$ to be the set of streams that terminate at node k; find a set β_{ij}, $i \in \mathcal{B}_k$, $j = 1 \ldots, P_k$ such that:

$$\sum_{i \in \mathcal{B}_k} \beta_{ij} b_i \leq V, \quad j = 1, \ldots, P_k \tag{3}$$

$$\sum_{j=1}^{P_k} \beta_{ij} = 1, \quad \beta_{ij} \text{ is binary}, \quad \forall i \in \mathcal{B}_k$$

Although in general these problems could be solved by linear integer programming, for the purposes of this paper we just implemented an exhaustive search, due to the relatively small search space in the cases evaluated.

5.3 Evaluation Results

We have simulated the network under the conditions described in section 5.1, and obtained both the session acceptance probability and the average number of hops, as a function of the offered load, for sessions composed of 10, 15 and 20 streams. We also obtained the same performance measures for an 8-node ShuffleNet, with 2 transmitters and 2 receivers per node (i.e., the same size as the WDM network under evaluation), using exactly the same sessions. Figure 2 shows the session acceptance probability for sessions composed of 15 streams, both for the WDM reconfigurable network, and for the ShuffleNet. In that figure, we also plot the upper bound. Figure 3 shows a comparison of the session acceptance probability for 10, 15 and 20 streams/session, both for the WDM network (using the routing and reconfiguration heuristic) and the ShuffleNet. As shown in Figure 3, for the same load, a session with a higher number of streams will have a higher probability of being accepted, because the bandwidth of the individual streams will be lower, thus allowing more freedom in arranging them. This is always true for the ShuffleNet. For the reconfigurable network, however, at very high loads, this trend is reversed - the performance for sessions with smaller number of streams is better because (for the sessions accepted) it is possible to dedicate a link to each stream. The main conclusions from these plots are:

- The fact that the session acceptance probability for the proposed heuristic is close to the upper bound indicates that there is no need to pursue the optimum solution, since the room for improvement is *at most* the difference between the two solutions.

- As expected, the reconfigurable network significantly outperforms the fixed-topology network (the ShuffleNet), even under uniformly-addressed traffic. For example, for a 90% session acceptance probability, the reconfigurable network can carry twice the load of the ShuffleNet, for 15-stream sessions.

Figure 4 shows the average path length as a function of the offered load, for 15 streams per session (the plots for 10 and 20 streams/session are similar), and it further confirms our observation that the performance of the reconfigurable network is better than the fixed-topology one; while the average path length for the ShuffleNet is around 2 hops, the path length for the reconfigurable network, even at high loads, is close to 1 hop, which is, of course, the minimum possible.

6 Evaluation of the WDM Network for a Dynamic Traffic Model

In the previous section, we evaluated the reconfiguration/routing algorithm using a static traffic model, i.e., the routing of a single session with a fixed number of streams on an idle network. Although this evaluation was useful to show that there is no motivation to pursue the optimum, it is not an indication of the actual performance of the algorithm in a realistic environment. In this section, we evaluate the WDM network and the routing/reconfiguration in a realistic environment, where sessions arrive, are routed (or dropped), and if accepted stay in the network for a certain period of time. We first describe the operation of the network in such a scenario, keeping in mind the traffic requirements, and then evaluate its performance, which we compare to that of a centralized switch.

6.1 Operation of the WDM Network in a Dynamic Environment

The main difference between a WDM network and a traditional mesh network is that the former is able to dynamically change its topology; with current optical components, this can be accomplished in sub-millisecond time. The ability to change the network topology at will during operation, in a de-centralized fashion, gives rise to the following issues:

Reconfiguring links in use: For stream traffic, if a link in use is reconfigured, the stream has to be interrupted for re-routing. This interruption corresponds to the time to "clear the network", reconfigure the links, and re-start service. Video/audio streams can tolerate interruptions and delay variations as long as the receiving end pre-buffers a certain amount of data, to keep playing during interruptions. For interactive traffic, however, the amount of pre-buffering cannot be very large as it adds latency to the communication; it is generally recognized that the latency for interactive communications should be less than 200 ms. The decision of *when* to reconfigure the network represents a tradeoff between performance (in terms of session blocking probability) and the number of times an established stream is re-routed during its lifetime. The extremes for this tradeoff are:

- Reconfigure only the idle links: this has the advantage that existing connections are never disturbed, but at high loads, it is unlikely that a link is completely idle; the network will find itself "locked" in a random configuration that is far from optimal, and will generally perform worse than a regular fixed-topology network under the same traffic conditions.
- Reconfigure the whole network at each arriving session: one could consider the arriving session and the sessions already established in the network as a new "larger" session to be routed on an empty network. This has the advantage that the network is always optimal, but a given stream might be re-routed an excessive number of times.

Control of the network: Control of the network is distributed. As long as the network remains strongly-connected[1], a small amount of bandwidth can be reserved in each link for management purposes. Therefore, reconfigurations that partition the network should not be allowed. In [16] we describe a simple modification of the Shortest Path with Reconfiguration algorithm to keep the network connected by performing a secondary reconfiguration, if necessary.

Considering the issues listed above, we propose the following model of operation for the reconfigurable network, which is illustrated in Figure 5:

- The network starts with some arbitrary strongly-connected topology (for control purposes).
- When a session arrives, it is either routed or blocked; blocked sessions are cleared.
- The reconfiguration/routing algorithm used to accept a session and route it is:

[1]There is at least one path between every pair of nodes in the network.

Step 1: Try to route the session on the current network topology, using the traditional shortest path algorithm. Prior to routing each stream in the session, temporarily prune from the network topology those links that do not have enough free bandwidth to support it. A session can be accepted only if the routing of all its component streams is successful.

Step 2: If the session was successfully routed step 1, accept this route. The streams already established in the network will not be disturbed. Otherwise, proceed to step 3.

Step 3: Since no path was found in step 1, the only alternative to accept this session would be to reconfigure existing connections. We use the heuristic presented in section 4 to compute the new topology and routes, considering as our "session" the existing streams and the new session being added.

Step 4: If the reconfiguration/routing in step 3 was successful, we accept the new session and implement the reconfiguration. Otherwise, we block the incoming session and the network topology remains unchanged.

In summary, we reconfigure the network only when the cost of not doing so is to block a session; this way, we have the same performance as if we were to reconfigure at every session, while reducing the number of times a given stream is re-routed.

6.2 Simulation Results

In this section, we present simulation results for the WDM network in a dynamic environment, where sessions arrive according to a Poisson process, and if not blocked, stay for an exponentially-distributed amount of time. The performance measures of interest are:

- Session blocking probability: probability that an arriving session cannot be routed and is blocked.

- Average time between successive reconfigurations.

- Average path change for re-routed streams; the path change is defined as the difference between the longest and the shortest paths experienced by the stream during its lifetime. This is a measure of the delay jitter introduced by the reconfiguration, which has to be taken into account when defining the receive buffer sizes for video and audio streams.

Figure 6 shows the average path length in the WDM network, as a function of the session arrival rate. At low arrival rates, the average path length is that of the initial network topology (2 hops in this case), because there is no need for reconfiguration. As the traffic increases, the network reconfigures more often and the path length decreases, using less resources and enabling the network to accept more sessions.

Figure 7 shows the average time between reconfigurations, as a function of the session arrival rate. Note the time is given as a multiple of the session duration, i.e., a value of 10 means that the average time between reconfigurations is 10 times the lifetime of a session; in the average, approximately only one out of 10 sessions will be reconfigured. The figure shows that reconfiguration is seldom employed, becoming more frequent only at very high traffic loads.

We also computed the average change in the path (difference between the maximum and the minimum paths during the lifetime of the stream), and found it to be under 0.6 hops in all cases. Moreover, streams using higher bandwidths are less likely to have their routes changed than streams using lower bandwidths, when the network reconfigures. In Figure 8 we plot the average path change as a function of both the requested bandwidth per stream and the session arrival rate. The figure shows that, at low loads, most streams are not reconfigured; as the load increases, the network will tend to reconfigure more the low-bandwidth streams, in favor of the high-bandwidth ones.

We note that the WDM network can be thought of as a "distributed switch", where the switching function is performed at the nodes and the "center" of the network is completely passive. However, the same function could be performed by a centralized switch (such as an ATM switch); routing is trivial (all paths are of the form source $\rightarrow$ switch $\rightarrow$ destination).

So, it is important to determine if, from a performance point of view, there is any advantage in using a WDM network.

The two network configurations, using a "distributed switch" (WDM) and a centralized switch, are shown in Figure 9. For the centralized switch, there are two pairs of optical transmitters/receivers, one to carry the traffic from the node to the switch, and the other to carry the traffic from the switch to the node. A WDM network with the same number of optical transmitters and receivers would have two transmitters and receivers per node. We claim that the complexity of the two networks depicted in Figure 9 is approximately the same; the optical transceivers are more complex in the WDM case, but the "center" of the network is passive. On the other hand, in the centralized switch scenario, all the complexity is moved to the center of the network, and the optical transceivers are simpler. To complete the evaluation scenario, we still have to choose the transmitter/receiver data rates in the WDM case (V_{WDM}) and in the centralized switch case (V_{SW}). Two scenarios are possible:

- Same data rate for all transmitters and receivers in both situations, $V_{SW} = V_{WDM}$. We will denote this switch as "switch 1". Note that the bandwidth out of each node of the WDM network in this case is twice that of the switch.

- Same output bandwidth for each node in both situations, $V_{SW} = 2V_{WDM}$, i.e., although the same number of transmitters and receivers is used in both networks, the ones connected to/from the switch will have twice the data rate as the ones in the WDM network. We will call this "switch 2".

Note that switch 2 will always perform better than the WDM network under any traffic scenario; its performance can be seen as an upper bound to what the WDM network can achieve. This is so because the necessary and sufficient condition for a stream to be accepted at switch 2 is that bandwidth must be available both at the link from the source to the network, and at the link from the network to the switch. For the WDM network, this condition is necessary, but not sufficient, as the stream might need to go through intermediate nodes which must have bandwidth available.

We have simulated the WDM network described above, together with switch 1 and switch 2. In all cases, the exact same requests are offered to the three networks. For the all the evaluations in this section, sessions are composed of a single stream.

Figure 10 shows the session blocking probability for the three networks, as a function of the session arrival rate. The plot shows that, as expected, the performance of the WDM network is lower than that of switch 2, but not significantly. The performance of switch 1, however, is much lower than that of the WDM network. In Figure 10, the average bandwidth requested per stream was set to 25% of the link bandwidth; we repeated the evaluation for other values of the stream bandwidth and found similar results.

In summary, we have shown that the performance of a WDM reconfigurable network is much superior to that of a centralized switch with roughly the same amount of optical hardware (switch 1 in the discussion), and is comparable to the performance of a switch using optical transmitters and receivers at twice the rate of the WDM counterparts (switch 2 in the discussion). The WDM network will introduce an additional delay jitter due to stream re-routing, but we have shown that, under reasonable traffic conditions, this re-routing represents a small effect, and will affect more the low-bandwidth streams than the higher-bandwidth ones. Moreover, if a specific stream cannot be re-routed for some reason, one just needs to mark the transmitters/receivers it is using as locked, and that stream will not be disturbed by network reconfiguration.

7 Conclusions

In this paper, we considered the problem of routing multimedia streams in a WDM reconfigurable network. We presented an exact solution for the optimum routing and reconfiguration problem, based on linear integer programming. Since this solution is complex, we also proposed an heuristic algorithm to find a sub-optimal solution, and derived an upper bound in performance. We showed that the heuristic solution is close to the upper bound, which obviates the need to pursue the more complex exact solution. Finally, we evaluated the performance of the WDM network in a realistic environment, and showed that it compares favorably with a centralized electronic switch of equivalent complexity.

Appendix

A An Integer Programming Solution to the Reconfiguration and Routing Problem

In this section, we show that the optimum routing/reconfiguration problem formulated in section 2 can be exactly solved by linear integer programming. Defining:

T	:	Number of streams in the session
$\{r_i\}$	:	Required bandwidth for stream i, $i = 1, \ldots, T$
$\{s_i\}$	:	Source node for stream i, $i = 1, \ldots, T$
$\{d_i\}$	:	Destination node for stream i, $i = 1, \ldots, T$
K	:	Total number of transmitters/receivers in the network
N	:	Number of nodes in the network
V	:	Bandwidth of each individual link
$\boldsymbol{P}$	:	Receiver distribution vector $(N \times 1)$; P_i is the number of receivers in node i
$\boldsymbol{L}$	:	Transmitter location matrix $(N \times K)$; $L_{ij} = 1$ if transmitter j is located in node i, otherwise $L_{ij} = 0$.
$\boldsymbol{R}$	:	Wavelength allocation matrix $(N \times K)$; $R_{ij} = 1$ if transmitter j is sending in a wavelength currently being received at node i, otherwise $R_{ij} = 0$.
$\boldsymbol{B}^i$	:	Destination vector $(N \times 1)$ for stream i, $i = 1, \ldots, T$; $B^i_{s_i} = 1$, $B^i_{d_i} = -1$, and $B^i_j = 0$ for $j = 1, \ldots, N$; $j \neq s_i$; $j \neq d_i$
$\boldsymbol{X}^i$	:	Routing vector $(K \times 1)$ for stream i; $X^i_j = 1$ if stream i is routed through transmitter j, $i = 1, \ldots, T$, $j = 1, \ldots, K$

The problem formulation can be expressed as:

GIVEN: $K, N, V, T, \boldsymbol{L}, \boldsymbol{P}; \{s_i\}, \{d_i\}, \{r_i\}, i = 1, \ldots, T$

MINIMIZE: Average path length

$$\sum_{i=1}^{T} r_i \sum_{j=0}^{K} X_j^i \tag{4}$$

WITH RESPECT TO: $\boldsymbol{R}$, $\boldsymbol{X}^i, i = 1, \ldots, T$

UNDER CONSTRAINTS:

1. Communication is one-to-one, i.e., there is only one transmitter and one receiver per wavelength. All receivers will be tuned to some wavelength to keep network connectivity, even if they are not being used for stream communication.

$$\sum_{i=1}^{N} R_{ij} = 1, \quad j = 1, \ldots, K \tag{5}$$

2. Node i has only P_i receivers.

$$\sum_{j=1}^{K} R_{ij} = P_i, \quad i = 1, \ldots, N \tag{6}$$

3. There should be a path from every source to every destination. This is equivalent to writing a set of flow conservation equations, for routing one unit of flow from the source to the destination of each stream in the session.

$$(\boldsymbol{L} - \boldsymbol{R})\boldsymbol{X}^i = \boldsymbol{B}^i, \quad i = 1, \ldots, T \tag{7}$$

4. The total bandwidth of the streams routed through a link should not exceed the link bandwidth:

$$\sum_{i=1}^{T} r_i X_j^i \leq V, \quad j = 1, \ldots, K \tag{8}$$

5. Integer constraints: receivers cannot be "divided".

$$\boldsymbol{R} \qquad \text{is binary} \tag{9}$$

No bifurcation of flow (in a packet-switched network, this condition can be relaxed, in which case the stream might be "divided" into several routes):

$$\boldsymbol{X} \qquad \text{is binary} \tag{10}$$

The objective function (4) and constraints (5), (6), (7) and (8) define a non-linear optimization problem; the objective function is linear, but the constraint set (more specifically, equations (7) - the flow conservation equations) is not. When constraints (9) and (10) are added, it becomes a non-linear integer optimization problem.

However, the non-linearity in the constraint set comes just from the $\boldsymbol{R}\boldsymbol{X}^i$ product in equation (7). By using the fact that $\boldsymbol{R}$ and $\boldsymbol{X}^i$ are binary variables, and by increasing the number of equations and free variables, we can convert the routing/reconfiguration problem into a *linear* integer programming problem. We add to the set of free variables the $N \times K$ binary matrices $\boldsymbol{Z}^i$, $i = 1, \ldots, T$, subject to the following new constraints:

$$Z^i_{jk} \leq X^i_k \tag{11}$$

$$Z^i_{jk} \leq R_{jk} \tag{12}$$

$$i = 1, \ldots, T; \quad j = 1, \ldots, N; \quad k = 1, \ldots, K$$

Equation (7) then becomes :

$$\boldsymbol{L}\boldsymbol{X}^i - \boldsymbol{Z}^i\boldsymbol{1} = \boldsymbol{B}^i \tag{13}$$

where $\boldsymbol{1}$ is a $K \times 1$ vector with 1 in all positions.

In summary, by adding $\boldsymbol{Z}^i$ to the list of free variables, replacing equation (7) with equation (13), and adding inequalities (11) and (12) to the constraint set, the reconfiguration/routing problem becomes a linear integer programming problem, which can be solved by standard techniques such as the branch-and-bound method [4]. It should be noted that, for a given fixed topology (i.e., given $\boldsymbol{R}$), this problem reduces to the well-known multicommodity flow problem.

B Formal Description of The Shortest Path with Reconfiguration Algorithm

INPUTS:

- A WDM network with N nodes; node i has S_i optical transmitters and P_i optical receivers, $i = 1, \ldots, N$. All transmitters and receivers are either *locked* or *free*; the locked transmitters and receivers are in a certain given topology, and the free ones are not connected at all.
- A source node s and a destination node d.

OUTPUT: The path from s to d with the minimum number of hops, using, if necessary, free transmitters and receivers.

ALGORITHM:

Step 1: Create a set of nodes P, initially empty, and a set of nodes T, initially containing all the nodes in the network. For each node in the network, associate a label l; initially, assign $l(s) = 0$, $l(i) = \infty$ for $i \neq s$. Create also two sets of nodes A and B, initially empty.

Step 2: Let i be the node with the smallest $l(i)$ (if there are many, choose one at random). If $l(i) = \infty$, there is no path between s and d using the locked receivers and transmitters; go to step 5. Otherwise, move i from the set T to the set P. If node i has a free transmitter and the set A is empty, also put node i in A.

Step 3: Update the label set as follows: for $j = 1, \ldots, N$, $j \neq i$, if there is a transmitter in node i connected to a receiver in node j, make $l(j) \leftarrow \min\{l(j), l(i) + 1\}$.

Step 4: If $i \neq d$, return to step 2; otherwise, the shortest path using the locked transmitters and receivers has been found; proceed to step 5.

Step 5: Create a set of nodes P', initially empty, and a set of nodes T', initially containing all the nodes in the network. For each node in the network, associate a label l'; initially, assign $l'(d) = 0$, $l'(i) = \infty$ for $i \neq d$.

Step 6: Let i be the node with the smallest $l'(i)$ (if there are many, choose one at random). If $l'(i) = \infty$, go to step 8. Otherwise, move i from the set T' to the set P'. If node i has a free receiver and the set B is empty, put node i in B and go to step 8. Also, if $i = s$, again go to step 8.

Step 7: Update the label set as follows: for $j = 1, \ldots, N$, $j \neq i$, if there is a receiver in node i connected to a transmitter in node j, make $l'(j) \leftarrow \min\{l'(j), l'(i) + 1\}$. Return to step 6.

Step 8: If the set A or the set B or both are empty, terminate. The shortest path is the one found in step 4, if any. Otherwise, let us denote by a the node in A and by b the node in B. Let $L_1 = l(d)$, $L_a = l(a)$ and $L_b = l'(b)$. If $L_1 \leq L_a + L_b + 1$, the path found in step 4 is the shortest. Otherwise, tune the free transmitter in a to the free receiver in b, and use the path from s to a, the $a - b$ link, and the path from b to d.

References

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

[2] D. Bertsekas and R. Gallager, *Data Networks*, Prentice Hall, New Jersey, 1987.

[3] M. S. Bazaraa, J. J. Jarvis and H. D. Sherali, *Linear Programming and Network Flows, 2nd ed.*, John Wiley & Sons, New York, 1990.

[4] F.S. Hillier and G.J. Lieberman, *Introduction to Operations Research*, 5th edition, New York, McGraw-Hill, 1990.

[5] B. Mukherjee, "WDM-based local lightwave networks. I. Single-hop systems," *IEEE NETWORK*, vol.6, no.3, May 1992, pp 12-27.

[6] B. Mukherjee, "WDM-based local lightwave networks. II. Multihop systems," *IEEE NETWORK*, vol.6, no.4, July 1992, pp 20-32.

[7] I. Habbab, M. Kavehrad, C. Sundberg, "Protocols for Very High-Speed Optical Fiber Local Area Networks Using a Passive Star Topology," *IEEE Journal on Lightwave Tech.*, vol. LT-5, no. 12, Dec. 1987, pp 1782-1793.

[8] N, Mehravari, "Improved Multiple Access Schemes for Very High-Speed Optical Fiber Local Area Networks Using a Passive Star Topology," *IEEE GLOBECOM 1989*, Dallas, Nov. 27-30, 1989.

[9] I. Chlamtac and A. Ganz, "Design Alternatives of Asynchronous WDM Star Networks," *IEEE ICC 1989*, Boston, June 11-14, 1989.

[10] A. Ganz and Z. Koren, "WDM Passive Star - Protocols and Performance Analysis," *IEEE INFOCOM 1991*, Bal Harbour, Florida, April 9-11, 1991.

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

[12] 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.

[13] J.-F. Labourdette and A. Acampora, "Logically Rearrangeable Multihop Lightwave Networks," *IEEE Trans. on Communications*, vol.39, no.8, pp 1223-30, Aug. 1991.

[14] J.-F. Labourdette and A. Acampora, "Partially Reconfigurable Multihop Lightwave Networks," *IEEE GLOBECOM 1990*, San Diego, Dec. 2-5, 1990.

[15] H. Crowder, E. L. Johnson, M. Padberg, "Solving Large-Scale Zero-One Linear Programming Problems," *Operations Research*, Vol. 31, No. 5, Sept-Oct 1983, pp 803-34.

[16] C. Noronha and F. Tobagi, "Routing of Streams in a WDM Reconfigurable Network," *Technical Report*, Computer Systems Laboratory, Stanford University (in preparation).

[17] A.S. Acampora, M.J. Karol, M.G. Hluchyj, "Terabit Lightwave Networks: The Multihop Approach," *AT&T Technical Journal*, vol. 66, issue 6, Nov/Dec 1987, pp 21-34.

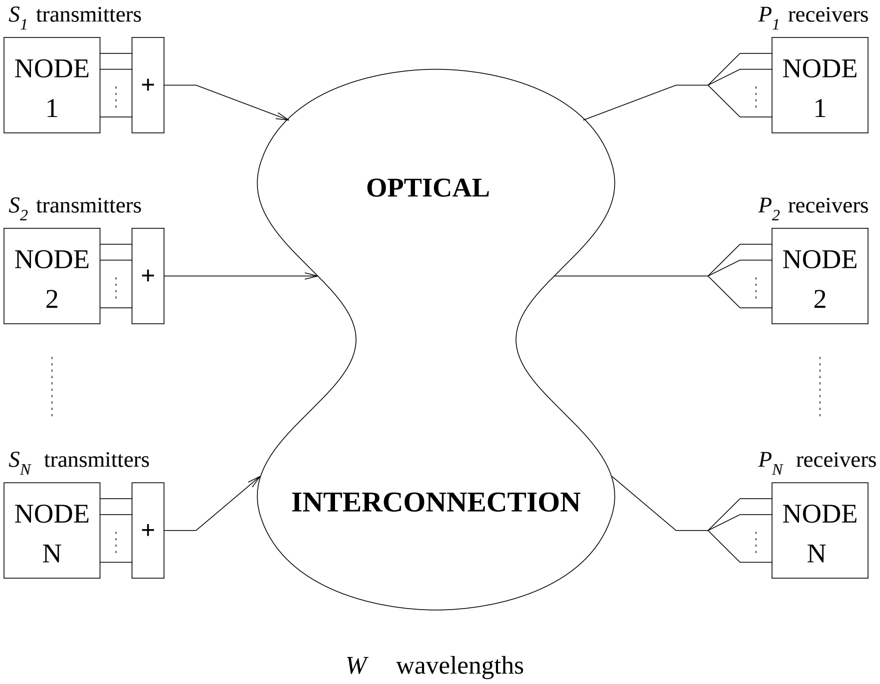

Figure 1: The WDM Network

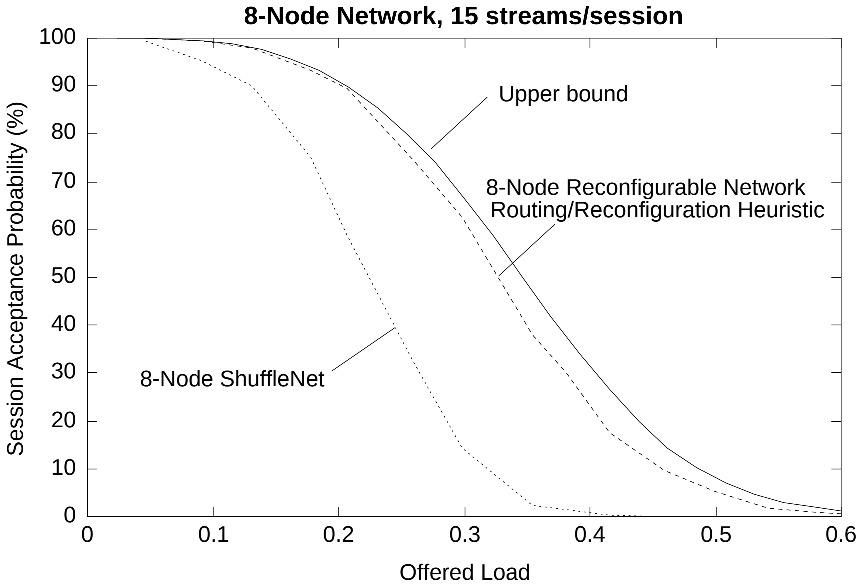

Figure 2: Session Acceptance Probability for 15 streams/session

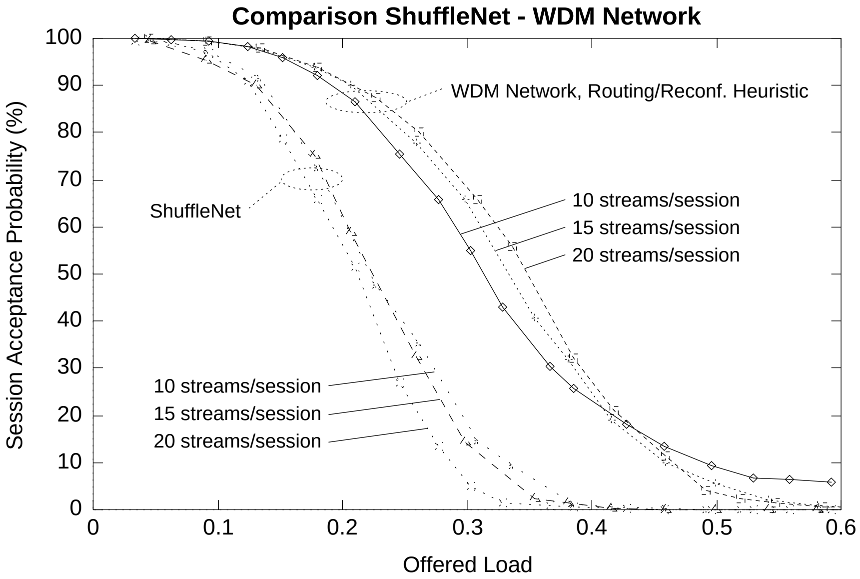

Figure 3: Comparison of the Session Acceptance Probability for the ShuffleNet and the WDM Network

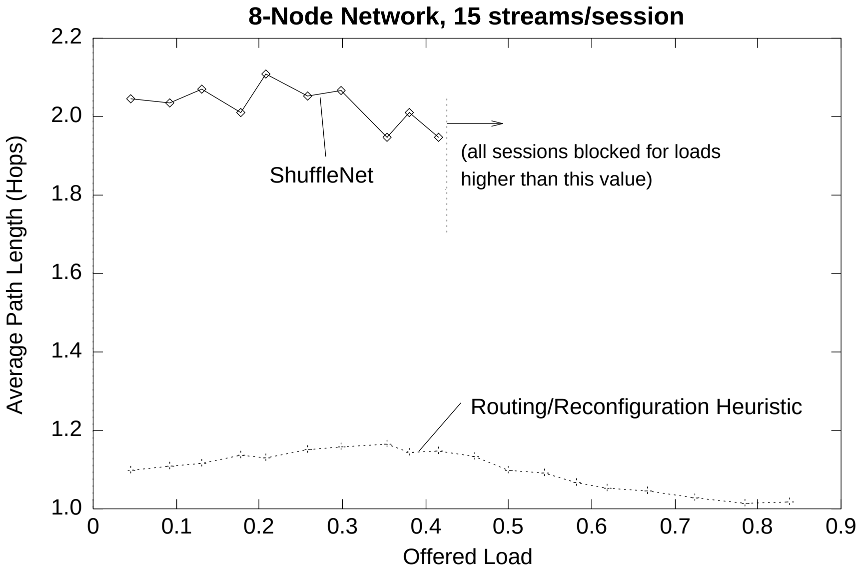

Figure 4: Average Path Length for 20 streams/session

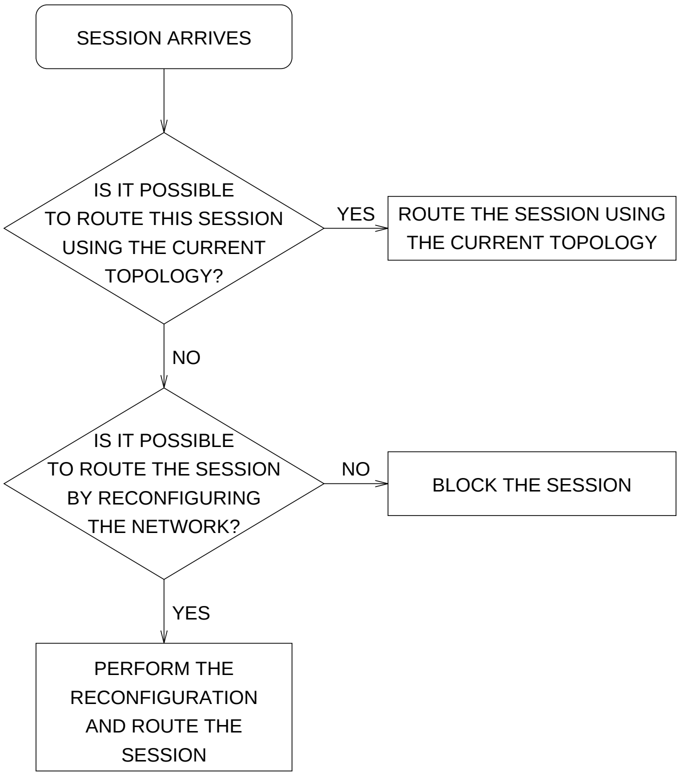

Figure 5: Dynamic operation of the WDM Reconfigurable Network

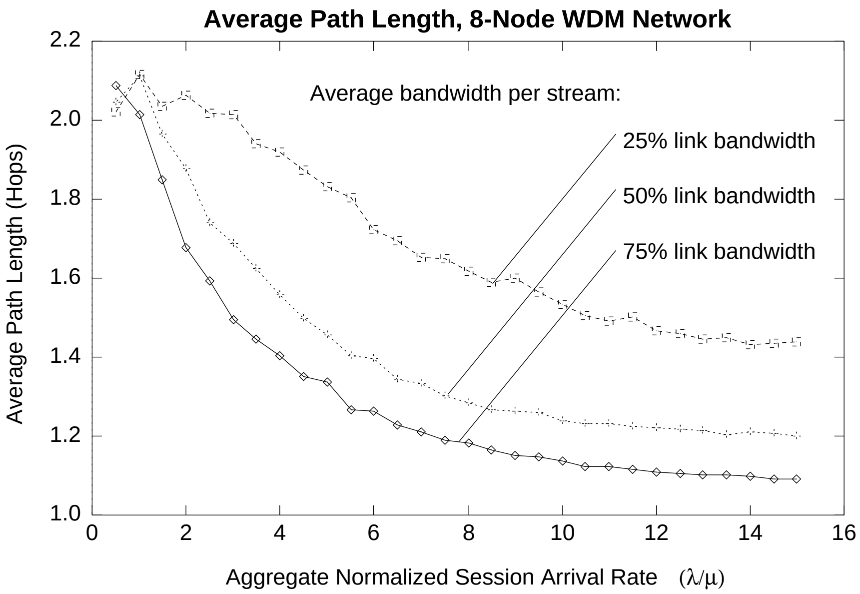

Figure 6: Average Path Length in the WDM Network

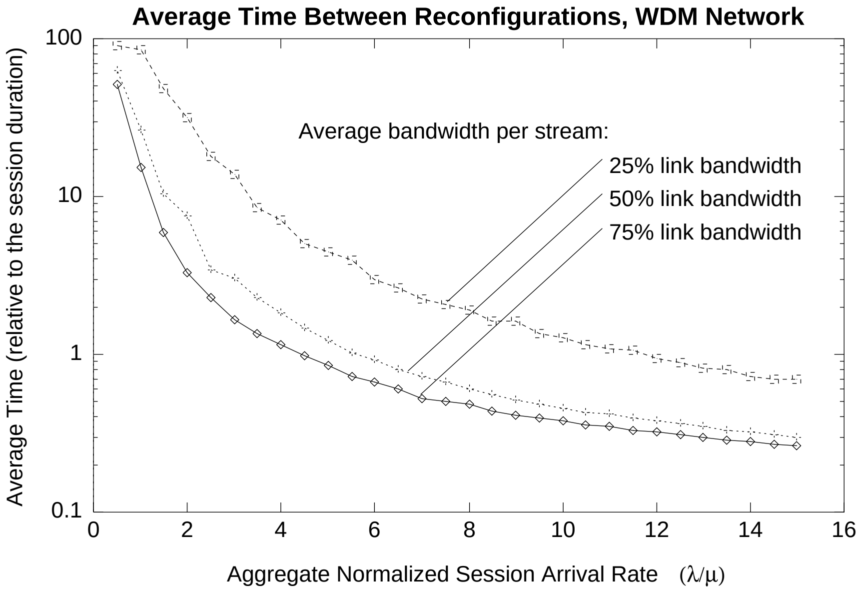

Figure 7: Average Time Between Reconfigurations in the WDM Network

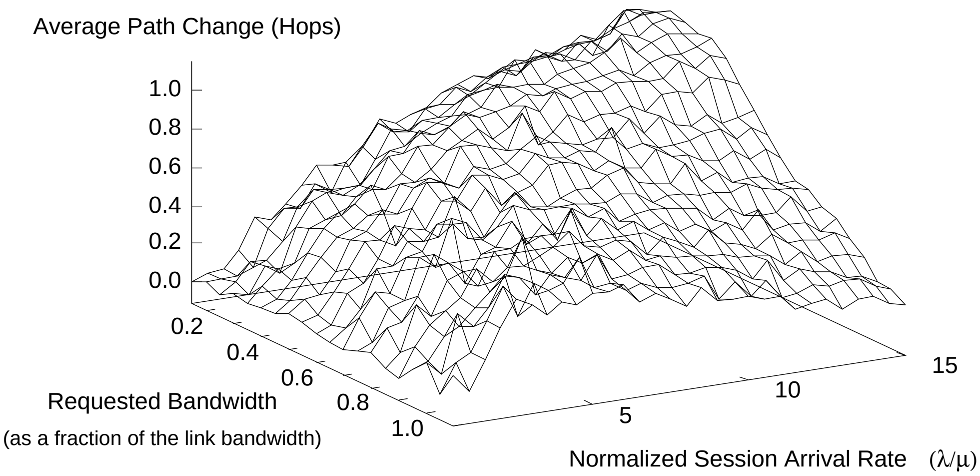

Figure 8: Average Path Change in the WDM Network

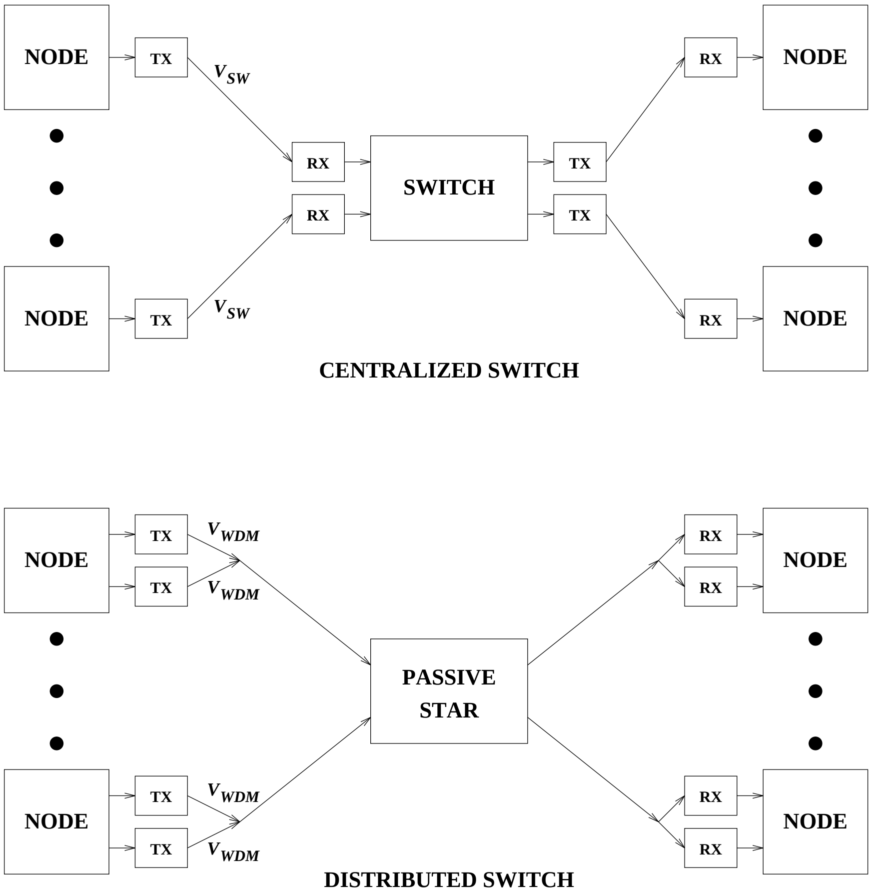

Figure 9: Distributed versus Centralized Switching

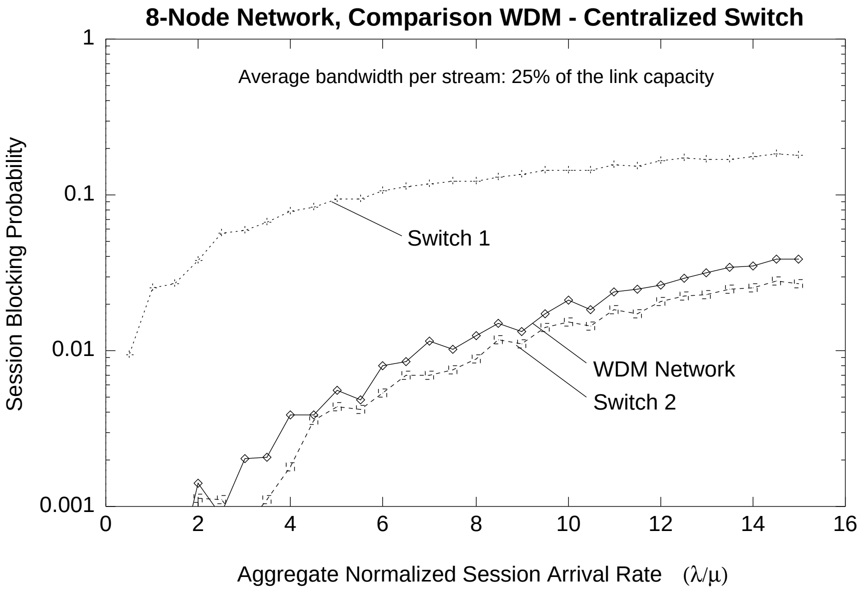

Figure 10: Blocking Probability results (single stream per session)