Bounds on the Blocking Performance of Allocation Policies in Wavelength Routing Networks and a Study of the Effects of Converters

Yuhong Zhu George N. Rouskas Harry G. Perros

TR-99-01

January 4, 1999

Abstract

We consider wavelength routing networks with and without wavelength converters, and several wavelength allocation policies. We show through numerical and simulation results that the blocking probabilities obtained through our previous analytical expressions for the random wavelength allocation and the circuit-switched case provide upper and lower bounds on the blocking probabilities for two wavelength allocation policies that are most likely to be used in practice, namely, most-used and first-fit allocation. Furthermore, we demonstrate that using these two policies has an effect on call blocking probabilities that is equivalent to employing converters at a number of nodes in the network. These results have been obtained for both single-path and general mesh topology networks, and for a wide range of loads. The main conclusion of our work is that the gains obtained by employing specialized and expensive hardware (namely, wavelength converters) can be realized cost-effectively by making more intelligent choices in software (namely, the wavelength allocation policy).

Keywords: Wavelength division multiplexing, wavelength routing networks, call blocking probability, wavelength allocation policies, wavelength converters

Department of Computer Science
North Carolina State University
Raleigh, NC 27695-7534

1 Introduction

Recent advances in wavelength division multiplexing (WDM) and optical switching make it possible to contemplate the deployment of wavelength routing networks that will provide backbone connectivity over wide-area distances and at very high data rates [9, 4]. A wavelength routing network consists of wavelength routers and the fiber links that interconnect them [7, 6, 8]. Wavelength routers are optical switches capable of routing a light signal at a given wavelength from any input port to any output port, making it possible to establish end-to-end lightpaths, direct optical connections without any intermediate electronics. The functionality of optical switches may be enhanced by employing wavelength converters, devices that are capable of shifting an incoming wavelength to a different outgoing wavelength [15]. Wavelength conversion is a desirable feature since it improves the performance of the network in terms of call blocking probability. However, any gains in performance must be weighted against the cost of wavelength converters.

While the operation of wavelength routing networks is expected to be similar to that of conventional circuit switched networks, several new issues arise which add significant complexity to the problems of design and performance evaluation of the former. Specifically, the existence of multiple distinct wavelengths makes it necessary to employ a wavelength allocation policy to assign one of the (possibly many) available wavelengths to an incoming call. Similarly, the wavelength conversion feature gives rise to new problems associated with evaluating the benefits of conversion and optimally placing the converters at the various network nodes. Also, dynamic (or adaptive) routing is tightly coupled with wavelength allocation, since it involves a search over available wavelengths in addition to a search over the possible paths for establishing a call.

The problem of computing call blocking probabilities under static (fixed or alternate) routing with random wavelength allocation and with or without wavelength converters has been studied in [1, 13, 2, 11, 16, 18]. The model presented in [1] is based on the assumption that wavelength use on each link is characterized by a fixed probability, independently of other wavelengths and links, and thus, it cannot capture the dynamic nature of traffic. In [13] it was assumed that statistics of link loads are mutually independent, an approximation that is not accurate for sparse network topologies. The work in [2] developed a Markov chain with state-dependent arrival rates to model call blocking in arbitrary mesh topologies and fixed routing; it was extended in [11] to alternate routing. While more accurate, this approach is computationally intensive and can only be applied to networks of small size in which paths have at most three links. A more tractable model was presented in [16] to recursively compute blocking probabilities assuming that the load on link i of a path depends only on the load of link $i-1$. Finally, a study of call blocking under non-Poisson input traffic was presented in [18], under the assumption that link loads are statistically independent.

Other wavelength allocation schemes, as well as dynamic routing are harder to analyze. First-fit wavelength allocation was studied using simulation in [3, 13], and it was shown to perform better than random allocation, while an analytical overflow model for first-fit allocation was developed in [12]. A dynamic routing algorithm that selects the least loaded path-wavelength pair was also studied in [12], and in [14] an unconstrained dynamic routing scheme with a number of wavelength allocation policies was evaluated. Except in [16, 17], all other studies assume that either all or none of the wavelength routers have wavelength conversion capabilities. The work in [16] takes a probabilistic approach in modeling wavelength conversion by introducing the converter density, which represents the probability that a node is capable of conversion independently of other nodes in the network. While this approach works well when the objective is the estimation of the expected call blocking performance, it cannot be used to calculate the actual blocking probability on individual paths when the placement of converters is known, nor can it be used to compare various converter placement strategies. Finally, in [17] a dynamic programming algorithm to determine the location of converters on a single path that minimizes average or maximum blocking probability was developed under the assumption of independent link loads.

Most of the approximate analytical techniques developed for computing blocking probabilities in wavelength routing networks [13, 2, 11, 18, 12, 14, 17] make the assumption that link blocking events are independent and amount to the well-known *link decomposition* approach [10], while the development of some techniques is based on the additional assumption that link loads are also independent. Link decomposition has been extensively used in conventional circuit switched networks where there is no requirement for the *same* wavelength to be used on successive links of the path taken by a call. The accuracy of these underlying approximations also depends on the traffic load, the network topology, and the routing and wavelength allocation schemes employed. While link decomposition techniques make it possible to study the qualitative behavior of wavelength routing networks, we believe that more accurate analytical tools are needed to both efficiently evaluate the performance of these networks, as well as to tackle complex network design problems, such as selecting the optical switches where to employ wavelength converters.

The authors have considered the problem of computing call blocking probabilities in mesh wavelength routing networks with fixed and alternate routing and random wavelength allocation in [19]. Unlike previous studies, we have developed iterative *path decomposition* algorithms [10] for analyzing arbitrary network topologies. Specifically, we analyze a given network by decomposing it into a number of path sub-systems. The latter are analyzed in isolation using our results on blocking in a single path in a wavelength routing network [20]. The individual solutions are appropriately combined to form a solution for the overall network, and the process repeats until the blocking probabilities converge. Our approach accounts for the correlation of both link loads and link blocking events, giving accurate results for a wide range of loads and network topologies. It also allows non-uniform traffic, in the sense that call request arrival rates can vary for each source-destination pair. Finally, our algorithms can compute call blocking probabilities in a mesh network where only a fixed but arbitrary subset of nodes are capable of wavelength conversion.

In this paper we study the blocking performance of several wavelength allocation policies for various network topologies and traffic patterns. Our main conclusions are as follows. First, we show that the most-used and first-fit policies have very similar performance for all calls in a network, regardless of the number of hops used by the calls. We also demonstrate that the random policy and the circuit-switched case (i.e., a system with converters at all nodes), for which analytical solutions exist for networks of large size, provide lower and upper bounds on the call blocking probability under the first-fit and most-used policies. We also present results which indicate that the effect of using these two policies is "equivalent" to using the random policy but employing a number of converters in the network. Overall, our results contradict previous

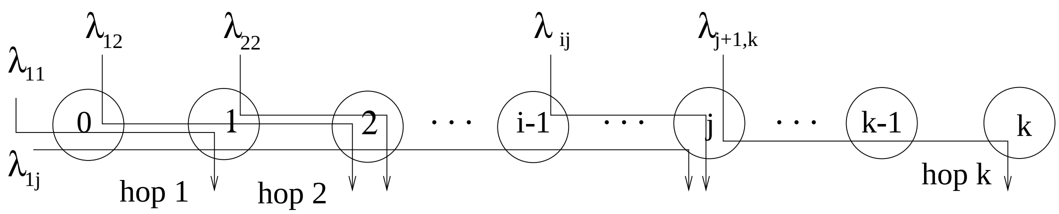

Figure 1: A k-hop path

studies which, concentrating only on random wavelength allocation, have suggested that sparse wavelength conversion is beneficial to wavelength routing networks. Although we identify regions of operation where converters do offer significant benefits, they are characterized by very high blocking probabilities, it is unlikely that networks will be designed to operate at these regions.

In Section 2 we study a single path in a wavelength routing network, and in 3 we consider both regular and irregular mesh network topologies. We conclude with a summary of our findings in Section 4.

2 A Single Path of A Wavelength Routing Network

Let us first consider a single path of a circuit-switched wavelength routing network, such as the k-hop path shown in Figure 1. A k-hop path consists of $k+1$ nodes labeled $0, 1, \cdots, k$, and hop $i, i = 1, \cdots, k$, represents the link between nodes $i-1$ and i. (Unless noted otherwise, the terms "hop" and "link" will be used interchangeably.) Each link in the path supports exactly W wavelengths, and each node is capable of transmitting and receiving on any of these W wavelengths. We let $\lambda_{ij}, j \geq i$, denote the Poisson arrival rate of calls that use hops i through j of the path, i.e., calls that originate at node $i-1$ and terminate at node j. For instance, λ_{22} is the arrival rate of calls that only use hop 2 (that is, those arriving at node 1 and leaving at node 2), while λ_{12} is the arrival rate of calls using hops 1 and 2 (refer to Figure 1). If the request can be satisfied, an optical circuit is established between the source and destination for the duration of the call. Call holding times are exponentially distributed with mean $1/\mu$. Also, $\rho_{ij} = \lambda_{ij}/\mu$ is the offered load of calls using hops i through j.

We define a "segment" of a k-hop path as a sub-path consisting of one or more consecutive links of the original path. We let n_{ij}, $j \geq i$, be a random variable representing the number of calls using hops i through j that are currently active in the path. We also let $f_{ij}, j \geq i$, be a random variable representing the number of wavelengths that are free on all hops i through j. As we shall shortly see, random variables n_{ij} and f_{ij} are part of the state description for the Markov process corresponding to the k-hop path.

In our model, we allow some of the nodes in the path to employ wavelength converters. These nodes can switch an incoming wavelength to an arbitrary outgoing wavelength. If no wavelength converters are employed in the path, a call can only be established if the *same* wavelength is free on all the links used by the call. This is known as the *wavelength continuity* requirement, and it increases the probability of blocking for calls using multiple hops. If a call cannot be established due to lack of available wavelengths, the call is blocked. On the other hand, if a call can be accommodated, it is assigned one of the wavelengths that are available on the links used by the call. If there multiple available wavelengths, a wavelength allocation policy must be employed to select a wavelength for the call. Different selection policies lead to different call blocking probabilities. In this paper we investigate the following wavelength allocation policies:

- *Random allocation:* a call is randomly assigned one of the wavelengths that are available on the links used by the call. This policy has been extensively studied in the literature, and we have developed approximate analytical algorithms to evaluate their performance [20, 19].
- *Most-used allocation:* the wavelength that is used on the largest number of links in the path (other, of course, than those used by the call) is assigned to the call; ties are broken arbitrarily. The objective of the policy is to keep more wavelengths available on long paths.
- *Least-used allocation:* the call is assigned the wavelength used in the smallest number of links in the path, with ties broken arbitrarily. Intuitively, this policy results to "wavelength fragmentation," and will lead to higher blocking probability for calls traveling over long paths.
- *First-fit allocation:* the wavelength are given labels in a fixed order, and the call is assigned the wavelength with the smallest label that is available on the links it uses. The objective of this allocation is to minimize wavelength fragmentation. As we shall show later, its performance is very close to that of the most-used policy, but it is easier to implement since there is no need to maintain information about global use of wavelengths.

In a path with wavelength converters, the corresponding allocation policy is used to assign a wavelength to the call within each segment of the path whose starting and ending nodes are equipped with converters. In addition to the above wavelength allocation policies, we will also study

- *Circuit-switched paths*: paths in which there are converters at all nodes. In circuit-switching, a call can be established as long as at least one wavelength (not necessarily the same one) is free on the links used by the call. Consequently, wavelength allocation is not an issue under a circuit-switching scenario, and all allocation policies, including the ones studied here, reduce to random allocation within each link.

In our study, we will use six different traffic load patterns to compare the wavelength allocation policies against each other and against circuit-switching. The six patterns are representative of the wide range of loading situations that one expects to encounter in practice. Figures 2 and 3 illustrate the six traffic patterns for a 10-hop path. Specifically, the figures plot the load ρ of each hop l, $l = 1, \cdots, 10$, in the path, defined as the sum of the offered loads $\rho_{ij}, i \leq l \leq j$, for all calls that use hop l, for each load pattern. In the "uniform" pattern, all hops are equally loaded. The "bowl" (respectively, "inverted bowl") pattern is such that the load decreases (resp., increases) from hop 1 to hop 5, and then it increases (resp., decreases) from hop 6 to hop 10. These patterns are shown in Figure 2. The "ascending" and "descending" patterns are such that the load increases or decreases, respectively, from hop 1 to hop 10. Finally, in the oscillating pattern the load at each hop alternates between a low and a high value. The last three load patterns are shown in Figure

3. Similar load patterns were used for shorter paths. To ensure that the results are comparable across the different patterns, the load values were chosen so that the total load (or, equivalently, the average load per hop) is the same for all patterns.

2.1 Policy Comparison for 2-hop Paths

We will first study the performance of the various wavelength allocation policies

2.1.1 Exact and Approximate Markov Processes

We have shown in [20] that the evolution of a 2-hop path with random wavelength allocation can be characterized by the four-dimensional Markov process $(n_{11}, n_{12}, n_{22}, f_{12})$. The first three random variables in the state description provide the number of active calls between the three source-destination pairs in the path, and the last random variable gives the number of wavelengths that are free on both links of the path. The state transition diagram of this Markov process is shown in Figure 5 for $W = 2$ wavelengths, and it is straightforward to see that the process is not time-reversible [20]. By modifying a few of the transition rates of this process, we have been able to derive an approximate time-reversible Markov process with the same state space, which has a product-form solution. If we let $G(W)$ denote the normalizing constant for a 2-hop path with W wavelengths per link, the solution of the approximate Markov process is given by [20]:

$$\pi_{random}(n_{11}, n_{12}, n_{22}, f_{12}) = \frac{1}{G(W)} \frac{\rho_{11}^{n_{11}} \rho_{12}^{n_{12}} \rho_{22}^{n_{22}}}{n_{11}! n_{12}! n_{22}!} \frac{\binom{f_{11}}{f_{12}} \binom{n_{11}}{f_{22} - f_{12}}}{\binom{n_{11} + f_{11}}{f_{22}}} \tag{1}$$

where $f_{11} = W - n_{11} - n_{12}$ and $f_{22} = W - n_{22} - n_{12}$. We have demonstrated in [20] that the blocking probabilities obtained through the product-form solution to the approximate Markov process are very close to the blocking probabilities obtained through the numerical solution to the original Markov process for a wide range of traffic loads.

Let us now consider the same 2-hop path with the most-used wavelength allocation policy. This policy can be modeled by a Markov process with the state description as that for the random policy: $(n_{11}, n_{12}, n_{22}, f_{12})$. The key difference is that, under the most-used policy, if, say, $n_{11} > n_{22}$, we know that there is at least one wavelength that is used on hop 1 but not used on hop 2. Thus, an incoming call that uses the second hop only will be assigned a wavelength that is already in use on the first hop, and will cause a transition to state $(n_{11}, n_{12}, n_{22} + 1, f_{12})$; similarly for $n_{22} > n_{11}$ and incoming calls using only the first hop. (Under the random wavelength allocation policy, the transition could be to either state $(n_{11}, n_{12}, n_{22} + 1, f_{12})$ or to state $(n_{11}, n_{12}, n_{22} + 1, f_{12} - 1)$ if the number of free wavelengths on both hops $f_{12} > 0$ and one of these wavelengths is assigned to the call.)

The state transition diagram of the Markov process for the most-used allocation policy is shown in Figure 6 for a 2-hop path with $W = 2$ wavelengths. Again, it is straightforward to verify that this Markov process is not time-reversible. Comparing to Figure 5, we note that despite having the same state space,

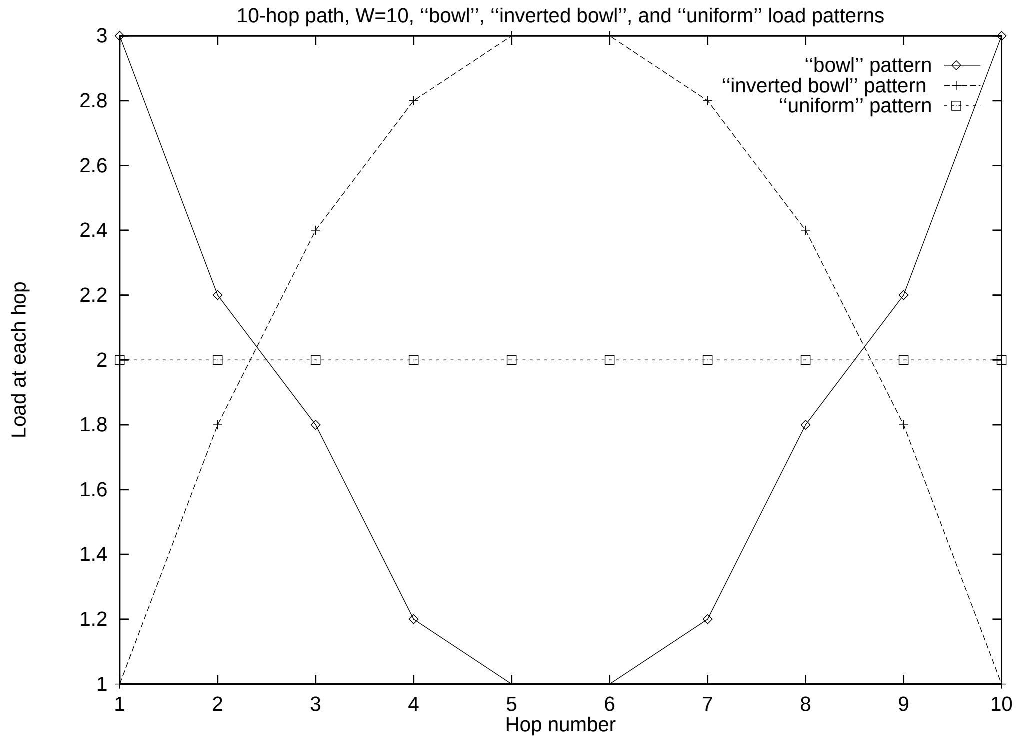

Figure 2: The "bowl", "inverted bowl", and "uniform" load patterns

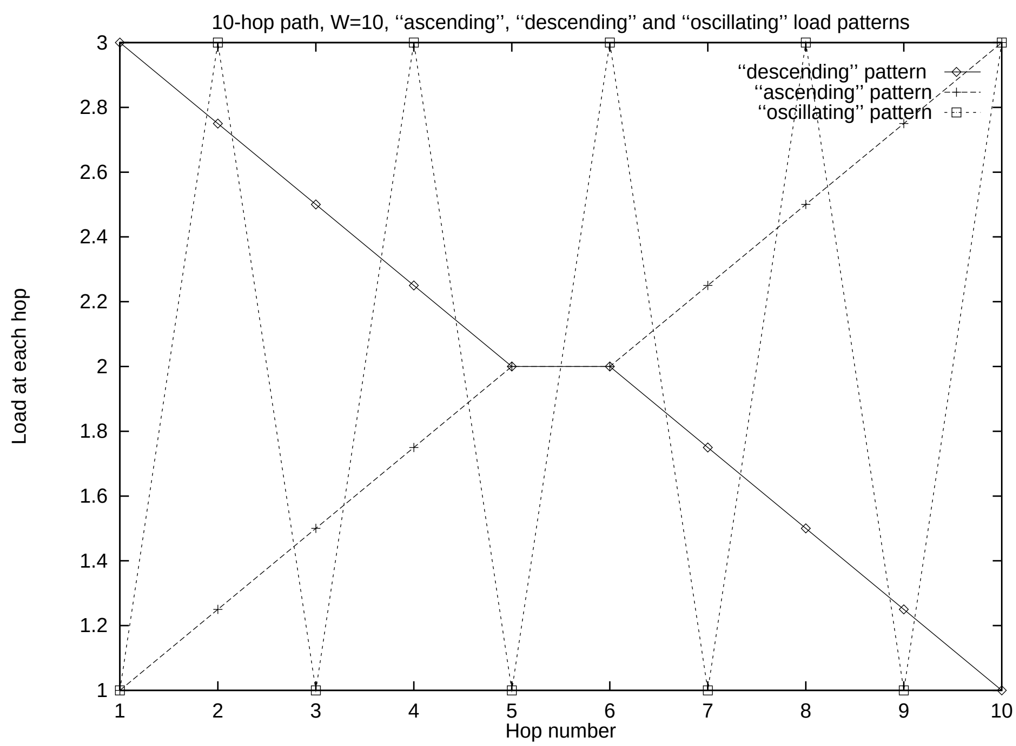

Figure 3: The "ascending", "descending", and "oscillating" load patterns

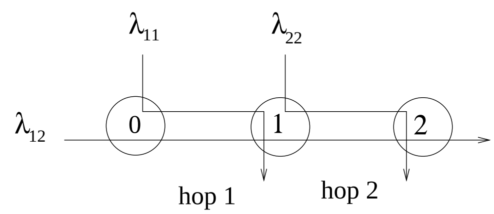

Figure 4: A 2-hop path

the two processes differ in two ways. First, some of the transition rates are different; for instance the transition rate from state (0,0,1,1) to state (1,0,1,1) is equal to $\lambda_{11}/2$ for the random allocation, but λ_{11} for the most-used allocation. Second, some of the transitions are missing in the new Markov process. For example, there is a transition from state (0,0,1,1) to state (1,0,1,0) under random allocation in Figure 5, but there is no such transition in Figure 6. Furthermore, since there is a transition from state (1,0,1,0) to state (0,0,1,1) in Figure 6, but no transition in the reverse direction, it is not possible to obtain an approximate time-reversible process by simply modifying some of the transition rates, as we were able to do for the random wavelength allocation policy. Although we do not have an approximate product-form solution for the most-used allocation policy, the state space for a 2-hop path is small enough that the solution to the Markov process can be obtained numerically for up to $W = 20$ wavelengths.

Based on similar arguments, it can be determined that the least-used wavelength allocation policy can also be modeled by a Markov process with the state description $(n_{11}, n_{12}, n_{22}, f_{12})$. The state transition diagram for this process is shown in Figure 7, and it can be easily verified that the process is not time-reversible.

If a converter is placed at node 1 of the 2-hop path in Figure 4 (the only interesting possibility in this case), the system becomes equivalent to a 2-hop circuit-switched path, and it can be described by the three-dimensional Markov process (n_{11}, n_{12}, n_{22}). Random variable f_{12} becomes redundant because calls continuing on both hops can now use *any* of the $(W - n_{12} - n_{22})$ available wavelengths on the second hop. It is well-known that this Markov process has the closed-form solution:

$$\pi_{cs}(n_{11}, n_{12}, n_{22}) \quad = \quad \frac{1}{G(W)} \frac{\rho_{11}^{n_{11}}}{n_{11}!} \frac{\rho_{12}^{n_{12}}}{n_{12}!} \frac{\rho_{22}^{n_{22}}}{n_{22}!} \tag{2}$$

It is also interesting to note that the normalizing constant in (2) is the same as that in (1). In Figure 8 we show the state space of a 2-hop circuit switched path with two wavelengths. Although this path is described by the above 3-dimensional Markov process, we include in the state description of Figure 8 the random variable f_{12} to make it easier to compare to Figures 5–7. For instance, the fact that there are no transitions into state (1,0,1,0) in the figure can be explained by recalling that $f_{12} = 0$ (i.e., that no wavelength is free on both links of the path) implies that calls traversing both hops are blocked. However, since exactly one wavelength is free on each hop (even if it is not the same one), calls traversing both hops cannot be blocked in the circuit-switched path, and the system will never enter state (1,0,1,0), but only state (1,0,1,1).

The first-fit wavelength allocation policy can be modeled by a Markov process with W state variables $(l_1, \cdots, l_W)$. Each random variable l_i corresponds to one of the W wavelengths, and can take one of five

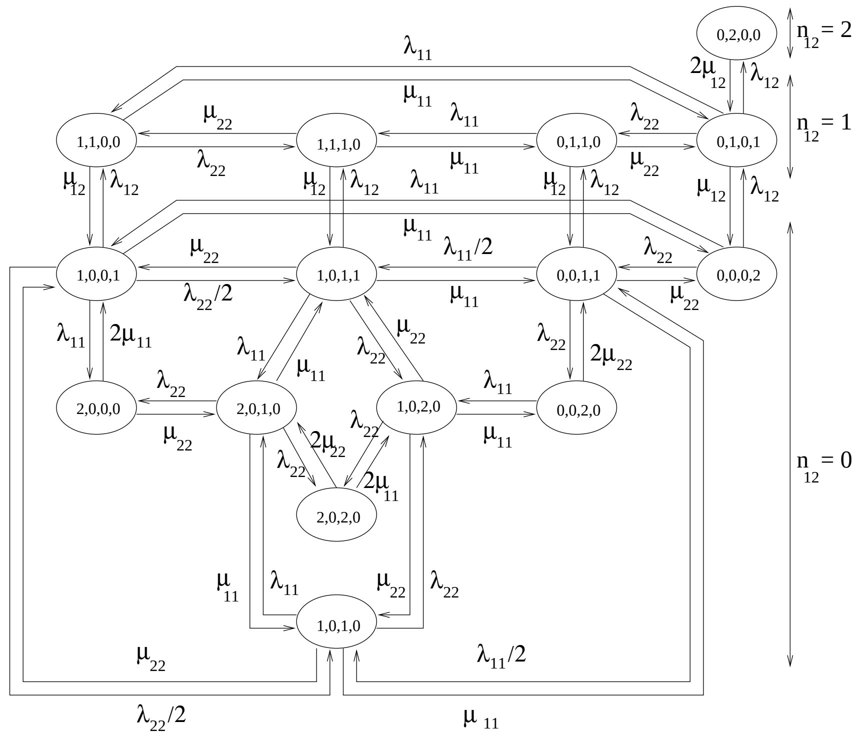

Figure 5: State space $(n_{11}, n_{12}, n_{22}, f_{12})$ of a 2-hop path with $W = 2$ wavelengths (random allocation)

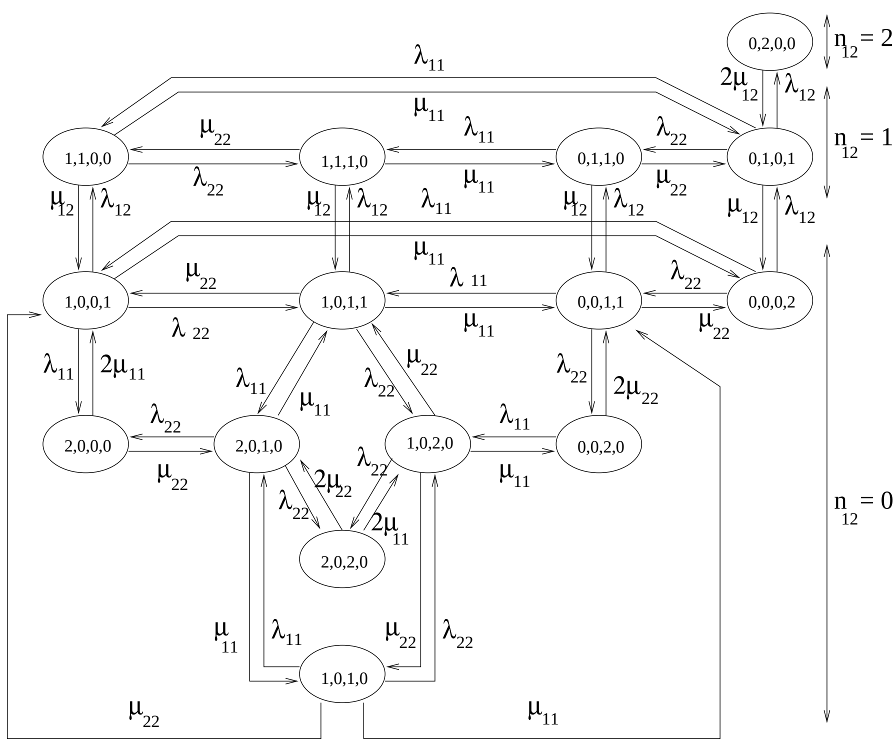

Figure 6: State space $(n_{11}, n_{12}, n_{22}, f_{12})$ of a 2-hop path with $W = 2$ wavelengths (most-used allocation)

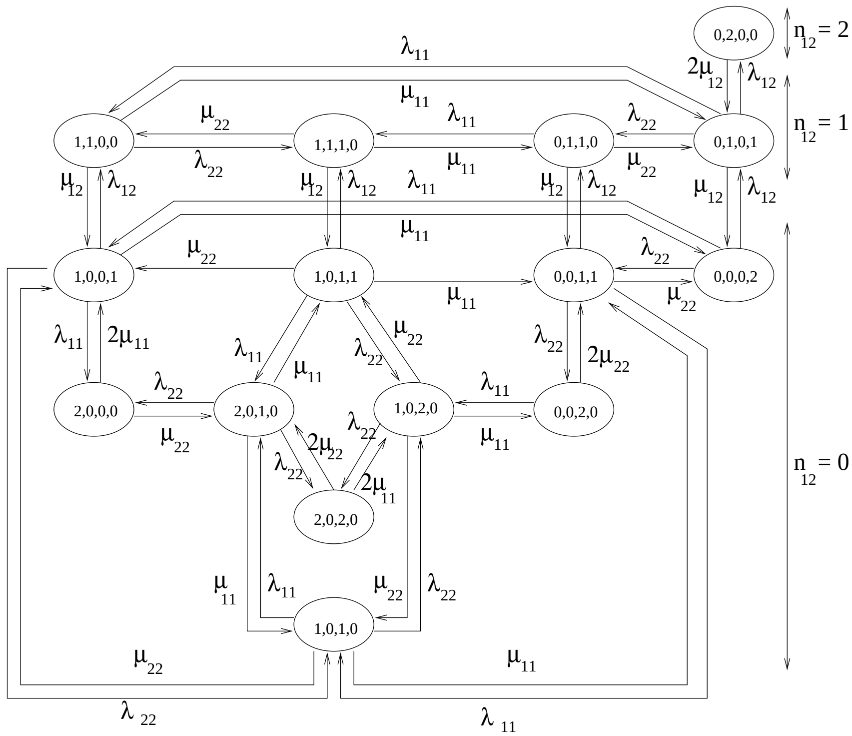

Figure 7: State space $(n_{11}, n_{12}, n_{22}, f_{12})$ of a 2-hop path with $W = 2$ wavelengths (least-used allocation)

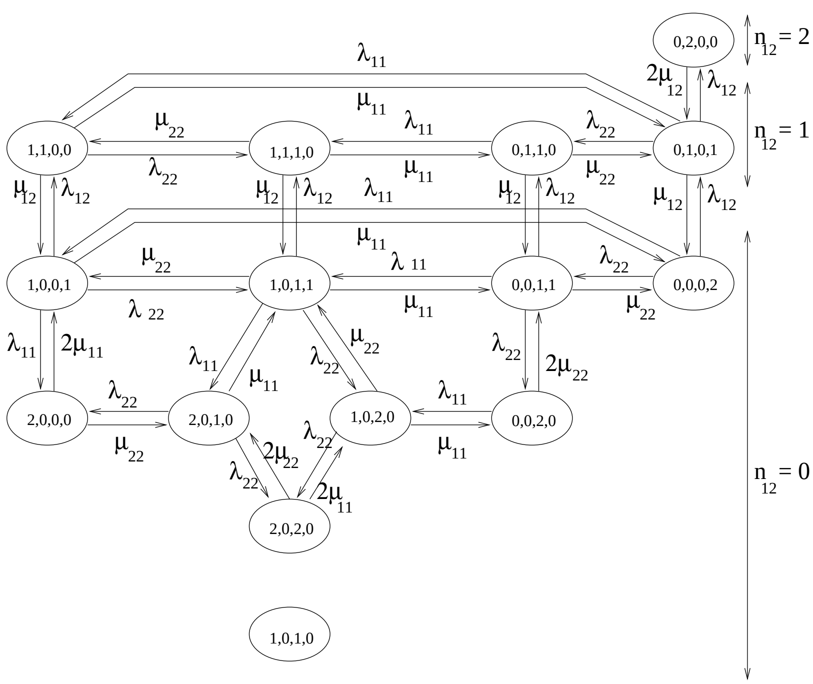

Figure 8: State space (n_{11}, n_{12}, n_{22}) of a 2-hop path with $W = 2$ wavelengths (circuit-switched)

values representing the status of wavelength i on the two links of the path: 0, if the wavelength is free on both links, 1, if it is free on the first link and busy on the second, 2, if it is busy on the first link and free on the second, 3, if the wavelength is used by two different calls on each link, and 4, if it is used by a call traversing both links of the path. The state space of this Markov process is quite different than that in Figures 5–8, and no direct comparisons can be made. Furthermore, the size of the state space is in the order of W^5, too large to obtain a numerical solution even for relatively small values of W. The main motivation for considering the first-fit policy in this work is that simulation results indicate that its performance is very similar to that of the most-used policy, while it is easier to implement.

2.1.2 Numerical Comparison

Let us first consider the blocking performance of the random, most-used, least-used, and circuit-switched systems for calls traversing both links of the 2-hop path. In Figures 5–8, the blocking states for these calls are those with the last random variable $f_{12} = 0$, i.e., those states in which neither of the two wavelengths is free on both links. We also observe that, except for state (1,0,1,0) at the bottom of the four figures, the transitions (and transition rates) in and out of all other blocking states are exactly the same for all four wavelength allocation policies. Consequently, we expect that the difference in the blocking probability experienced by calls traversing both links of the path under the different policies will be mainly due to the steady-state probability of blocking state (1,0,1,0).

Referring to Figure 8, we note that the corresponding Markov process never enters state (1,0,1,0). Thus, we expect that calls traversing both hops will experience the least blocking probability in a circuit-switched path. In Figure 6 (most-used policy) we note that there are two transitions into state (1,0,1,0), and four transitions out of it. The blocking probability will be higher under this policy compared to the circuit-switched case. The Markov process in Figure 5 (random policy) has two additional transitions into state (1,0,1,0) from states (0,0,1,1) and (1,0,0,1) with rates $\lambda_{11}/2$ and $\lambda_{22}/2$, respectively. Therefore, the blocking probability of these calls with the random policy will be higher than with the most-used policy. Finally, the Markov process in Figure 7 (least-used policy) has the same transitions as the one in Figure 5, but the transition rates into state (1,0,1,0) from states (0,0,1,1) and (1,0,0,1) are λ_{11} and λ_{22}, respectively. Therefore, we expect that these calls will experience the highest blocking probability under the least-used policy.

We now note that the lower the blocking probability for calls traversing both hops, the larger the number of such calls accepted, and the larger the number of wavelengths they occupy, leaving fewer wavelengths available for calls using a single link (either the first or the second) of the path. Hence, we expect that the behavior of the four policies in terms of the blocking probability of calls using a single link of the path will be exactly the opposite of what was discussed above. Specifically, we expect the least-used policy to provide the lowest blocking probability for these calls, followed by the random, the most-used, and the circuit-switched policies, in that order.

The above conclusions, derived by direct comparison of the Markov processes, are in agreement with intuition. We have confirmed these conclusions by numerically comparing the blocking probabilities of the various policies for 128 different load values. Figures 9 and 10 show results for two cases corresponding to a uniform and descending load pattern, respectively, similar to the corresponding patterns shown in Figures 2 and 3, and for $W = 10$ wavelengths. More specifically, the arrival rates (refer also to Figure 4) used in Figure 9 were $\lambda_{11} = 0.?, \lambda_{12} = 0.?, \lambda_{22} = 0.?$, while for the results in Figure 10 we used $\lambda_{11} = 2.0, \lambda_{12} = 2.0, \lambda_{22} = 3.0$. The two figures plot the blocking probability for the three types of calls, namely, calls using the first hop only (label "hop 1" in the x-axis of the figures), calls using the second hop only (label "hop 2"), and calls using both hops (label "both hops"). We first note that the results are consistent with the corresponding traffic pattern. For instance, under uniform loading (Figure 9), calls using the first hop only experience the same blocking probability as hops using the second hop only, while in the descending pattern (Figure 10), due to the lower load offered to the second hop, the latter calls experience much lower blocking probability for all four policies. More importantly, the relative values of the blocking probabilities for the four policies are also consistent with our discussion above. Very similar results have been obtained for all 128 different load values that we have studied.

Finally, in Figure 11 we compare the most-used and first-fit policies for the same arrival rates as those used for Figure 10. As before, the blocking probabilities of the most-used policy were obtained through a numerical solution to the corresponding Markov process, while the values for the first-fit policy were obtained through simulation. We observe that the blocking probabilities values of the first-fit policy are almost identical to those of the most-used policy for all three types of calls. This result can be explained by noting that both policies attempt to maximize the number of wavelengths that are available for calls that use both hops of the 2-hop path by reducing the "fragmentation" of the set of wavelengths. The most-used policy assigns to an incoming call that uses a single hop of the path a wavelength that is already used on the other hop, if such a wavelength exists. On the other hand, the first-fit policy attempts to achieve the same goal by searching the set of wavelengths in a fixed order, thus increasing the chances that a wavelength used on a single hop will be assigned to an incoming call using the other hop. As we can see from Figure 11, the most-used policy is slightly more successful, but overall the blocking probability values of the two policies are very close. Similar results have been obtained for all 128 traffic loads that we have studied.

2.2 Policy Comparison for Longer Paths

Consider a k-hop path, $k > 2$, with the random wavelength allocation policy. Paths consisting of four links or less can be analyzed by solving the corresponding approximate time-reversible Markov process. For instance, a 3-hop path can be modeled by the 9-dimensional Markov process $(n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33}, f_{12}, f_{13}, f_{23})$ whose solution can be written down as a straightforward generalization of expression (1). Paths longer than four hops are analyzed using the iterative decomposition algorithm in [20] to obtain the call blocking probabilities. The analytical techniques developed in [20] are both accurate and efficient, and can be used when the path employs converters whose location is fixed but arbitrary. When all nodes in a k-hop path employ converters (the circuit-switched case), the call blocking probabilities can also be obtained by solving the corresponding k-dimensional Markov process through a straightforward generalization of expression (2). For very large k, when the computation of the normalizing constant becomes computationally expensive, a decomposition algorithm similar to the one in [20] can be used.

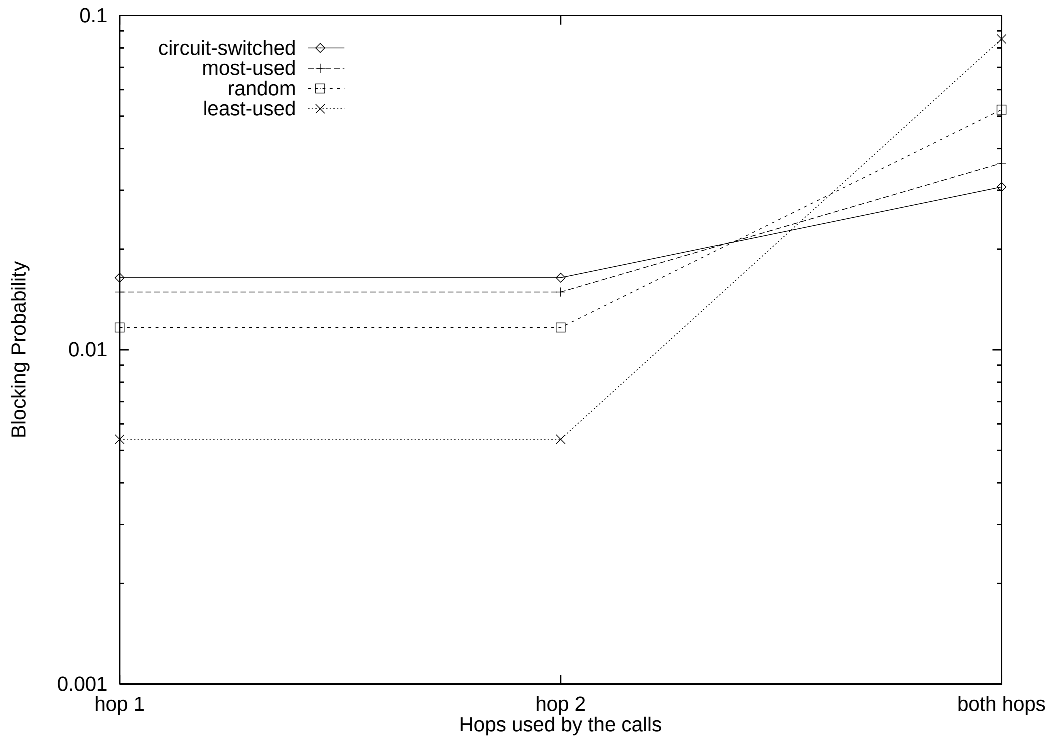

Figure 9: Policy comparison, 2-hop path, uniform traffic pattern

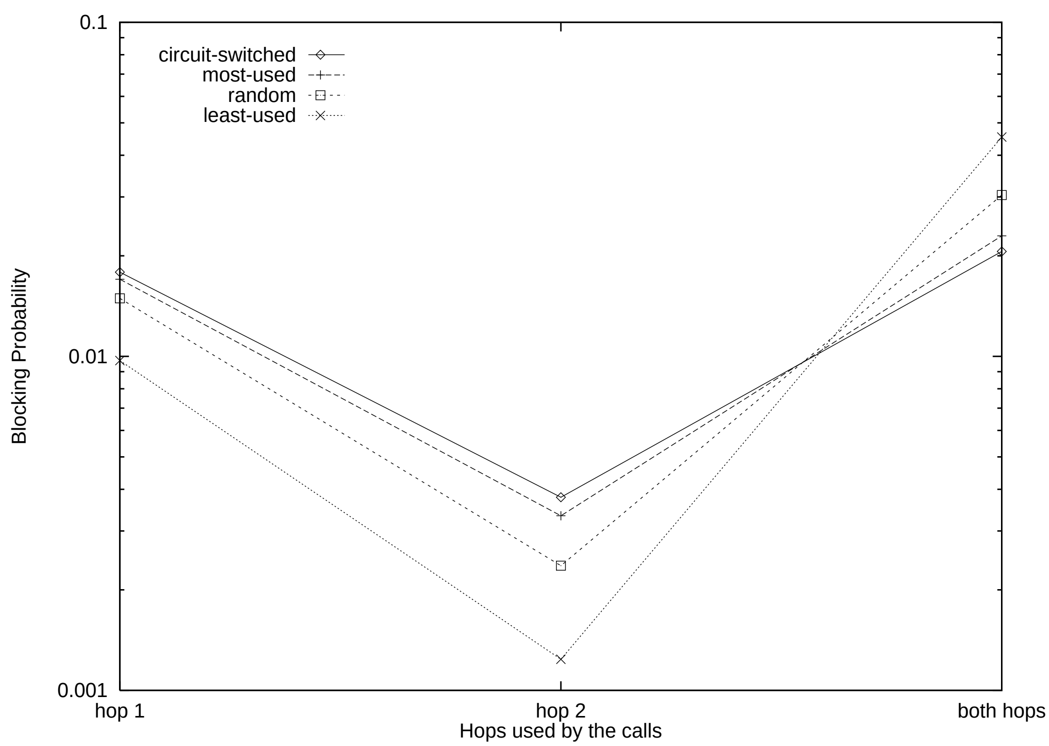

Figure 10: Policy comparison, 2-hop path, descending traffic pattern

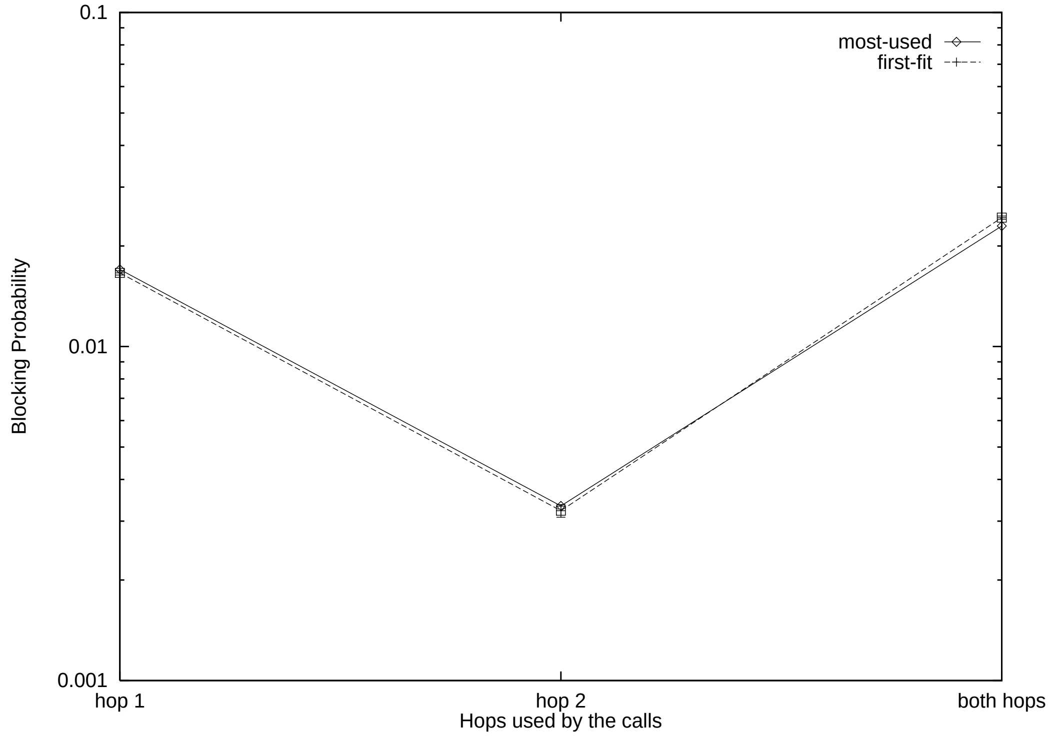

Figure 11: Most-used vs. first-fit allocation, 2-hop path, descending traffic pattern

Let us now consider the most-used wavelength allocation policy. It is not difficult to derive an exact Markov process to model a k-hop path, $k > 2$. However, unlike the 2-hop path case, the state description of this process for a k-hop path, $k > 2$, contains a number of random variables that is larger than the k^2 random variables in the state description of the process for the same path with random allocation. For a 3-hop path, for example, the state of the Markov process for most-used allocation is described by 12 random variables, the nine used to describe the corresponding process for random allocation plus three more. The larger state space, combined with the fact that an approximate closed-form solution is not possible for this process, makes it difficult to numerically obtain the call blocking probabilities under this policy for paths longer than two hops by directly solving the exact Markov process. Furthermore, developing an iterative algorithm for analyzing long paths by decomposing them into 2-hop path sub-systems which can be solved in isolation, similar to the algorithm developed for the random policy in [20], has turned out to be a difficult task. For such an algorithm, it is crucial to have an accurate estimate of the blocking probability due to the wavelength continuity requirement for calls traversing more than one sub-system. Since the different wavelengths are not equally utilized, as under random allocation, it is difficult to derive an approximate expression for blocking due to the wavelength continuity requirement that is accurate for a wide range of loads.

Using similar reasoning, it can be seen that developing approximate analytical techniques for the least-used and first-fit policies is also a rather hard task. Instead, we have used simulation to obtain the call blocking probabilities for paths with the most-used, least-used, and first-fit wavelength allocation policies [1].

2.2.1 Numerical Comparisons

In this section we present results for 6-hop and 10-hop paths, since the length of these paths (in hops) is representative of future backbone wavelength routing networks. Figures 12-15 correspond to a 6-hop path, while Figures 16-20 are for a 10-hop path.

In Figures 12 and 13 we compare the blocking probabilities for the four policies (random, most-used, least-used, and circuit-switched) and for the uniform and bowl traffic patterns, respectively. The two figures plot the blocking probability for the twenty one (21) different types of calls in a 6-hop path, numbered 1 through 21 in the x-axis. It is important to emphasize that the calls have been numbered such that numbers 1 through 6 correspond to the calls traversing a single hop in the path (i.e., hops 1 through 6, respectively), numbers 7 to 12 correspond to calls that traverse exactly two hops in the path, and so on.

From Figures 12 and 13 we observe that the relative behavior of the four policies is similar to that shown in Figures 9 and 10 despite the fact that the traffic patterns in the various figures are very different. Specifically, for calls using only a single hop (calls 1 through six in the figures), the least-used policy provides the lowest blocking probability, followed by the random policy, the most-used policy, and the circuit-switched case. However, for calls traversing multiple hops, the situation is reversed. The same behavior has been observed for other traffic patterns and paths of different length. We also note that, under the least-used policy, the blocking probability of calls using multiple hops increases significantly, and that the average blocking probability over all calls is higher than other policies. Therefore, we will not consider the least-used policy any further.

In Figure 14 we compare the most-used and first-fit policies for a 6-hop path with an inverted bowl traffic pattern. Again, as in Figure 11, we find that the two policies exhibit almost identical blocking performance, not only for the end-to-end call but for all calls, regardless of the number of hops used by the calls; very similar results have been obtained for all traffic patterns studied. Therefore, for the rest of the paper we will concentrate on the first-fit policy, since its implementation does not require that the network nodes maintain information about the global use of wavelengths. Specifically, our objective is two-fold. First, we will show that the values of the blocking probabilities obtained with this policy are bounded by the blocking probability values obtained by the random policy without converters and the circuit-switched case (equivalently, the random policy with converters at all nodes of a path). Second, we will demonstrate that, for calls traversing multiple hops, the gain (in terms of a reduction in blocking probability) obtained by employing the most-used over the random policy is roughly equivalent to using the random policy and deploying converters in the path or network.

In Figure 15 we compare the blocking probabilities of a 6-hop path obtained by the first-fit policy to

[1] Approximate analytical techniques based on overflow traffic have been developed for the first-fit policy in [12, 14]. However, these techniques are limited in that link blocking events are taken to be independent, an assumption that is not accurate over a wide range of traffic loads.

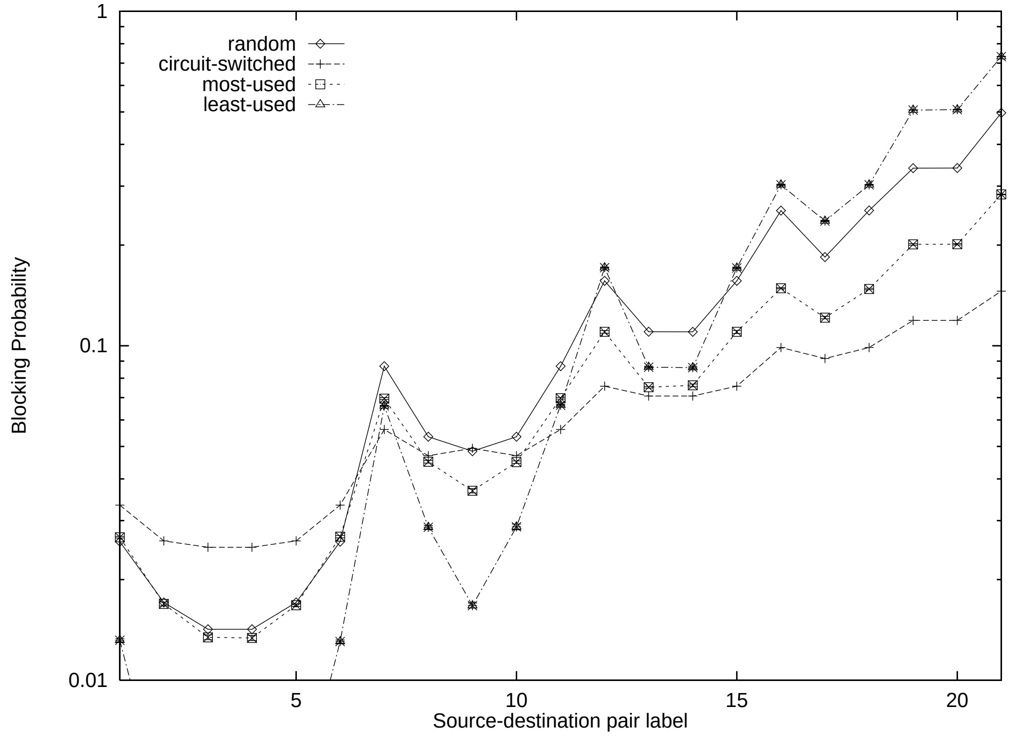

Figure 12: Policy comparison, 6-hop path, uniform traffic pattern

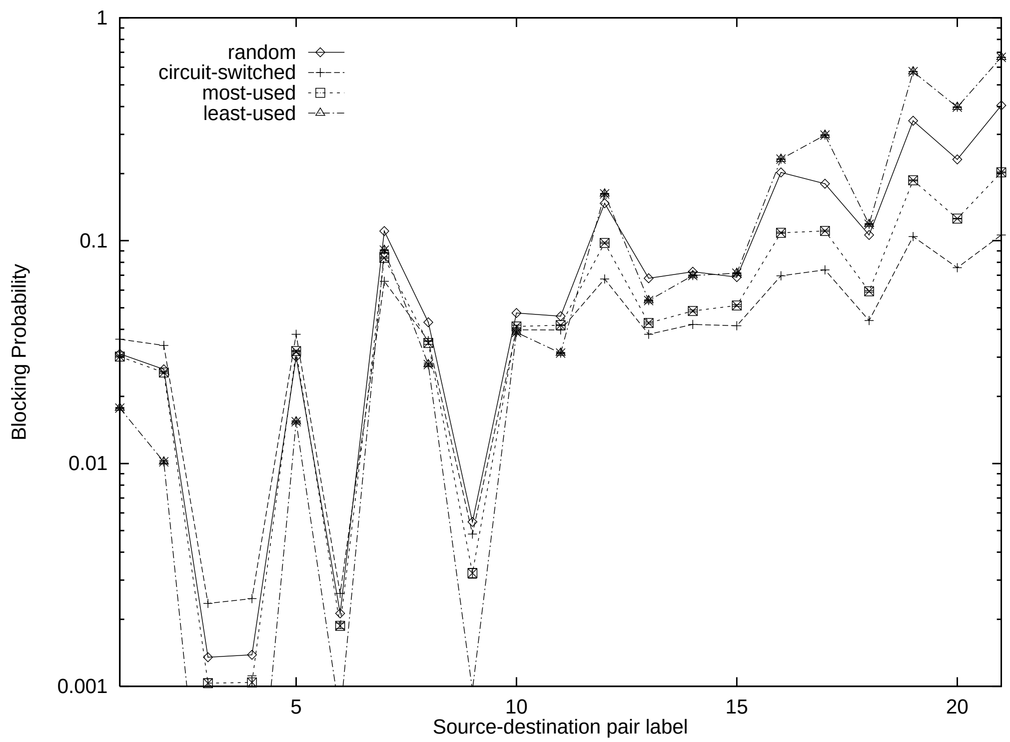

Figure 13: Policy comparison, 6-hop path, bowl traffic pattern

three other systems: a path with the random policy and no converters, a path with the random policy and one converter at node ?, and a circuit-switched path (i.e., a path with the random policy and converters at all nodes). As we can see, the first-fit policy has an effect similar to that of using the random policy and employing a converter in the path. This is a general result that has been observed for a wide range of traffic loads, and will be discussed in more detail shortly.

Figures 16- 20 present results for a 10-hop path and various traffic patterns. In Figures 16- 18 we compare the first-fit policy to the random (no converters) and circuit-switched cases for the inverter bowl, oscillating, and ascending traffic patterns, respectively. Two interesting observations can be made from the three figures. First, the blocking probability values of the first-fit policy are always between the corresponding values of the random and circuit-switched cases. In other words, the blocking probability values under the random and circuit-switched cases provide lower and upper bounds for the blocking performance of the first-fit policy [2]. Second, it appears that the first-fit policy is quite effective for reducing the blocking probability of calls traveling over multiple hops (which are the ones experience the highest blocking probability under the random policy) close to the level of the circuit-switched case.

In Figure 19 we compare the first-fit to the most-used policy for the descending traffic pattern. As before, the blocking probability values of the two policies match for all types of calls. Finally, in Figure 20 we attempt to quantify the effect of the first-fit policy in terms of "number of converters." Specifically, we plot the blocking probability values for a random policy with either three or five converters. Note that, by employing converters at some of the nodes, the blocking probability of calls traversing multiple hops improves, since converters reduce the requirement that the same wavelength be used on all hops of the path taken by the call. However, this improvement is at the expense of calls using a single hop, which now experience higher blocking probability. As we can see, the effect of using the first-fit policy in place of the random policy has an effect similar to employing converters.

3 Mesh Wavelength Routing Networks

In this section we further investigate the effects of the most-used and first-fit policies on the call blocking probabilities compared to the effect of converters, by studying two network topologies: a regular 5×5 torus network, and the NSFNET irregular topology.

3.1 The 5×5 Torus Network

Let us first consider the 5×5 torus network shown in Figure 21, with $W = 10$ wavelengths per link. Since there are 600 source-destination pairs in this network, it is impossible to present numerical results for all of them. We have therefore decided to present the call blocking probabilities for only 24 different source-destination pairs, and to summarize the results for the remaining pairs. The source-destination pairs for

[2]Note that the random policy provides a *lower* bound for calls using a single hop, and an *upper* bound for calls traversing multiple hops; the reverse is true for the circuit-switched path.

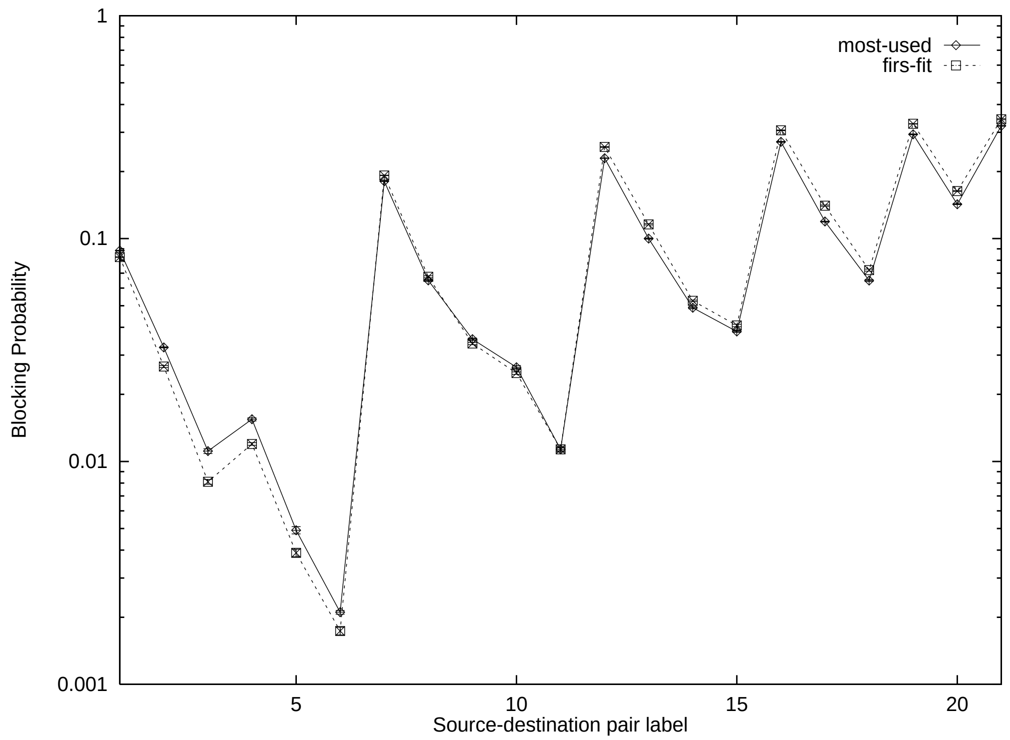

Figure 14: Most-used vs. first-fit allocation, 6-hop path, descending traffic pattern

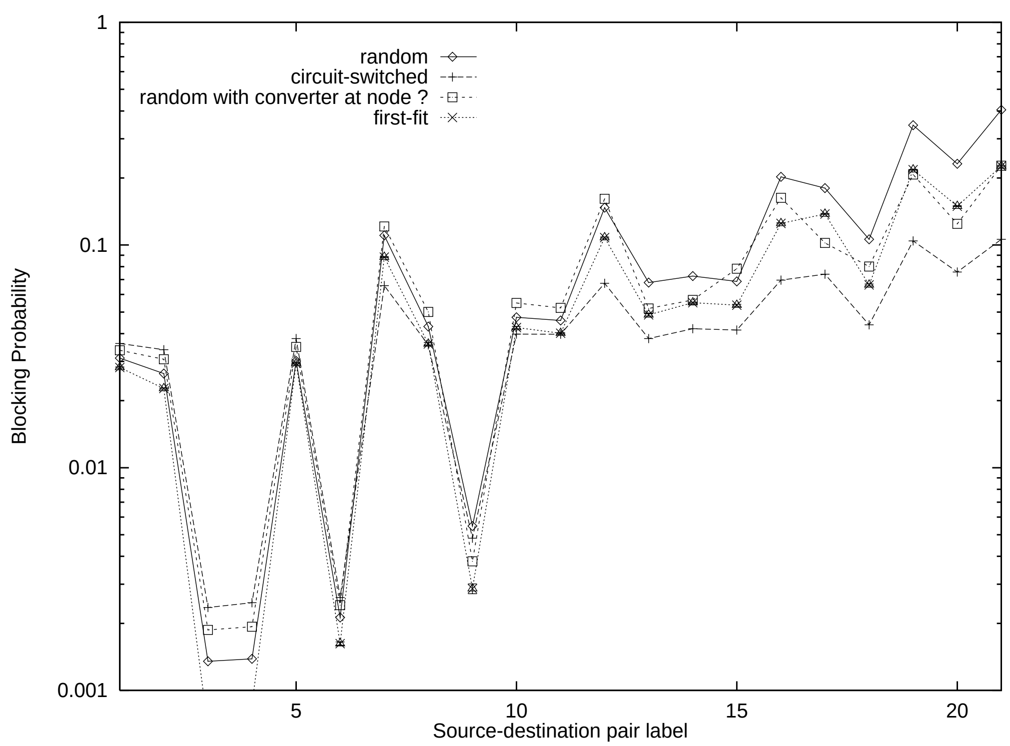

Figure 15: First-fit policy vs. random policy with converters, 6-hop path, bowl traffic pattern

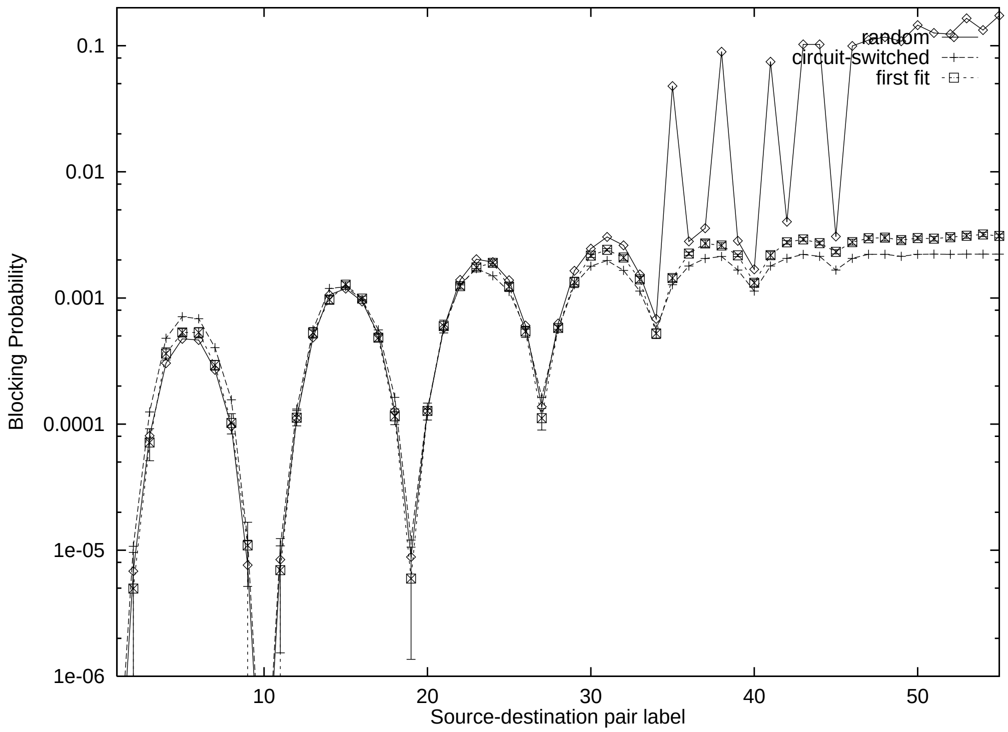

Figure 16: Policy comparison, 10-hop path, inverted bowl traffic pattern

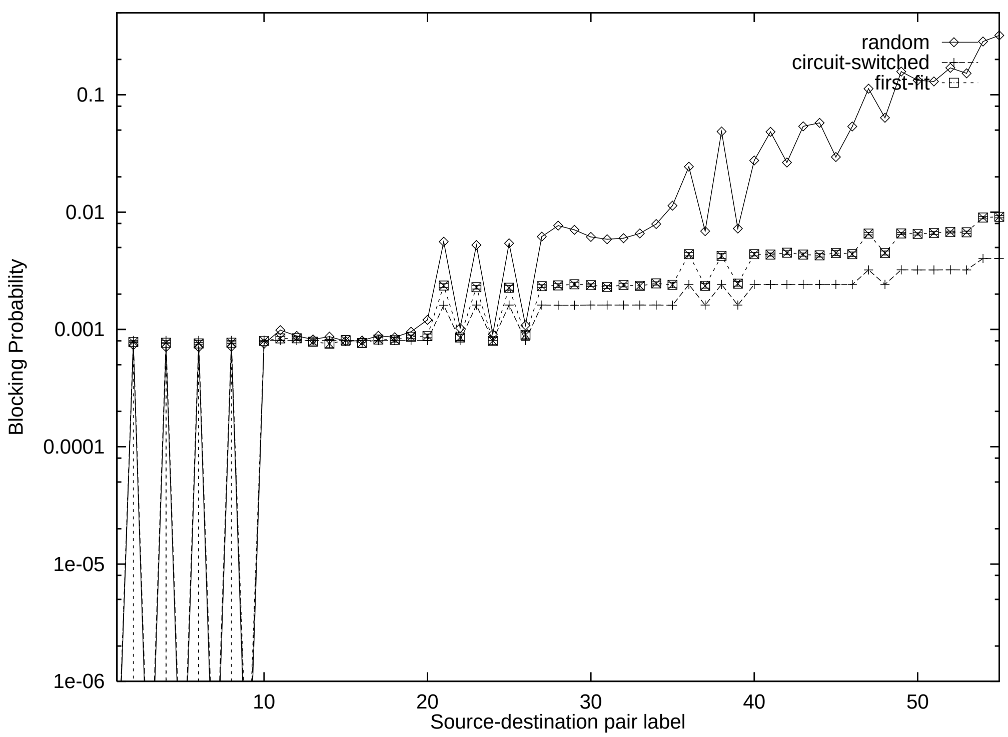

Figure 17: Policy comparison, 10-hop path, oscillating traffic pattern

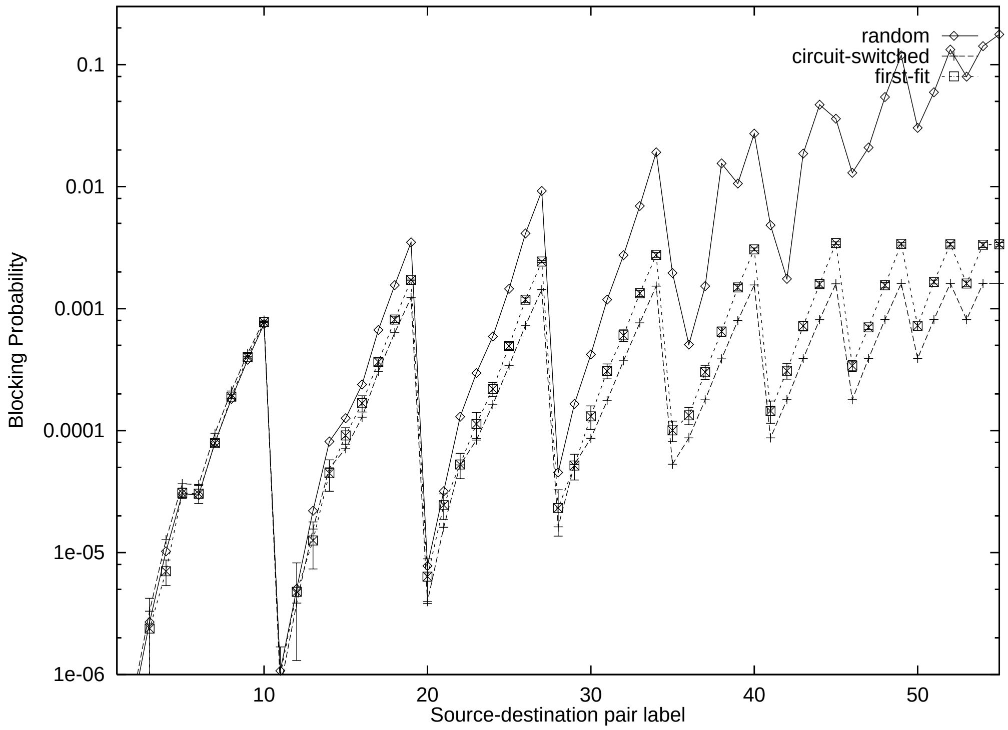

Figure 18: Policy comparison, 10-hop path, ascending traffic pattern

which detailed results are provided are those with node 1 as the source. Because of the regular topology, the selected pairs are a representative sample of the various source-destination pairs. Similar to previous figures, the source-destination pairs have been labeled such that numbers 1 through 4 correspond to pairs for which a 1-hop path is used, numbers 5 to 12 correspond to pairs for which a 2-hop path is used, and so on.

In our study we have used two traffic patterns. For the first pattern, the call arrival rates were selected such that

$$\lambda_{sd} = \begin{cases} 0.4, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 1 \\ 0.3, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 2 \\ 0.2, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 3 \\ 0.1, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 4 \end{cases} \quad (3)$$

This selection of arrival rates was intended to capture the locality of traffic that has been observed in many networks. The utilization of each link in the network for these arrival rates is in the range [3.140, 3.144]. The tight range of link utilizations can be explained by the fact that both the topology and the traffic load are symmetric. For the second pattern (which we will refer to as the random pattern), each arrival rate λ_{sd} was selected from a uniform distribution in the range (0.1,0.4).

In Figure 22 we compare the blocking probabilities obtained through the most-used and first-fit policies for the pattern based on locality of traffic. From the figure, we observe that calls using a single hop (labels 1 to 4 in the figure) experience the lowest blocking probability, calls traveling over two hops have the next

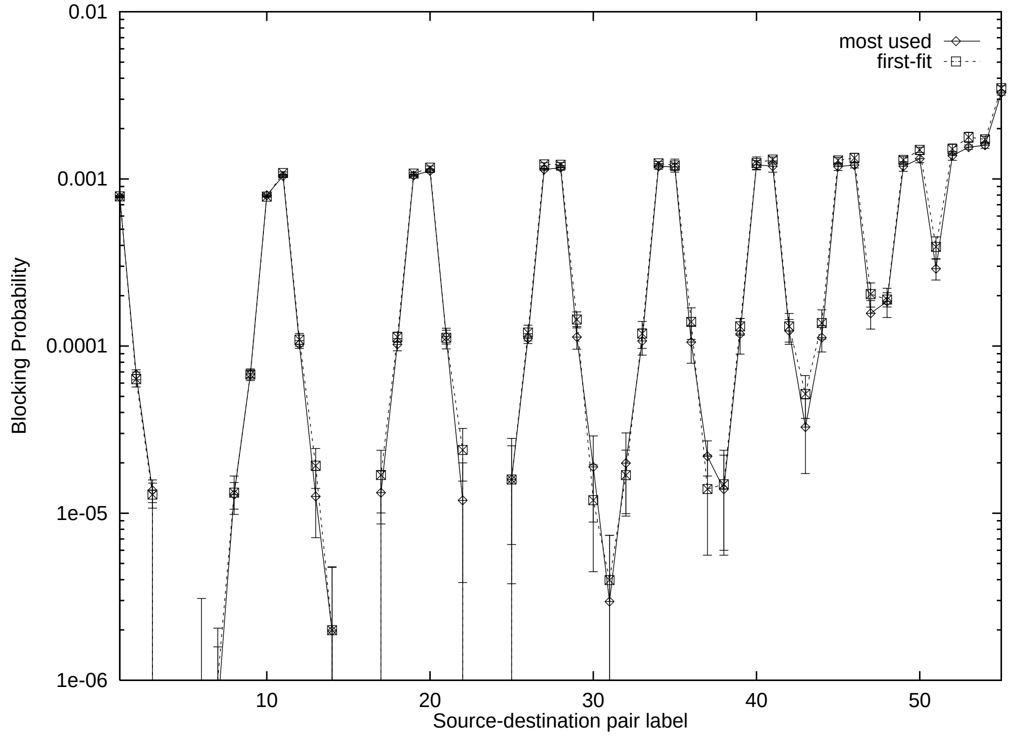

Figure 19: Most-used vs. first-fit allocation, 10-hop path, descending traffic pattern

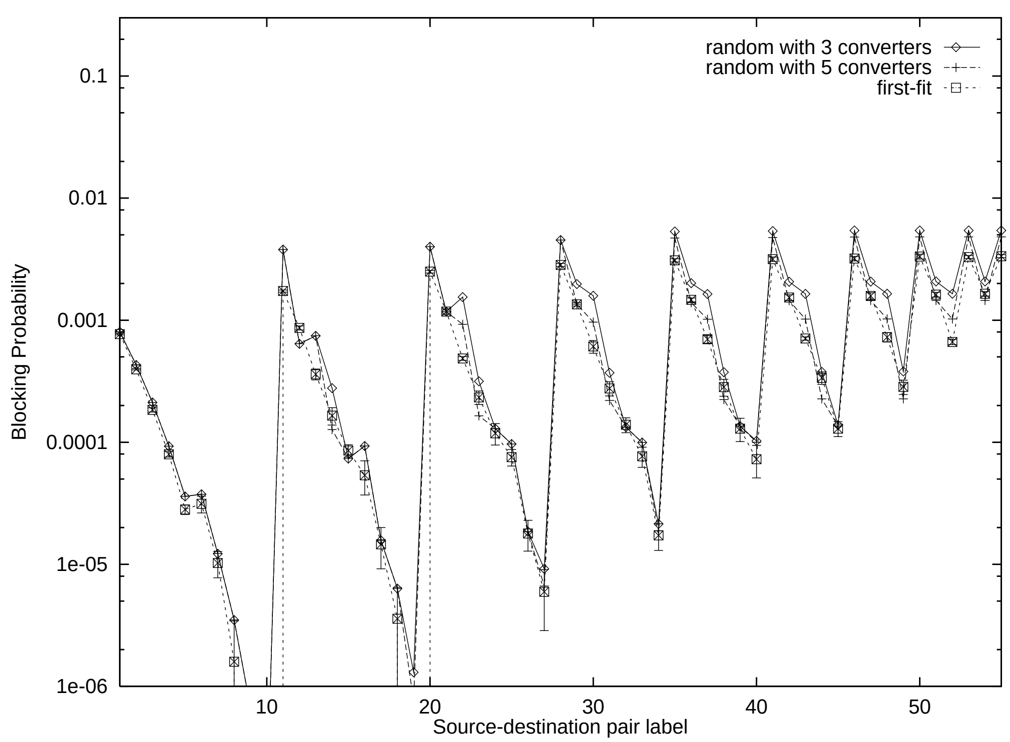

Figure 20: First-fit policy vs. random policy with converters, 6-hop path, descending traffic pattern

lowest blocking probability, and so on. (The fact that the blocking probability values are almost the same for all calls using the same number of hops is due to the symmetry of both the topology and of the traffic pattern.) We also see that the two policies result in almost identical blocking probability values for all calls, further confirming our claim that the (simpler) first-fit policy can be used as a quite accurate approximation of the most-used policy. Similar results, not shown here, have been obtained for the random traffic pattern.

In Figures 23 and 24, we compare the first-fit policy to the random policy and the circuit-switched case, under the two traffic patterns. It is clear from both figures that the blocking probability values under the first-fit policy are between those of the other two cases. However, there is also an important difference in the two figures. While the effect of the first-fit policy appears to have a significant effect for the traffic pattern based on locality (Figure 23), in that the blocking probability values for calls using multiple hops drops significantly from the corresponding values under the random policy, this effect is less pronounced in Figure 24 for the random traffic pattern. This behavior can be explained by noting that the blocking probabilities for calls 13 and higher in Figure 24 are more than 10%, about an order of magnitude higher than the values for the corresponding calls in Figure 23. At such high values, not many wavelengths are available for these calls, thus, the actual wavelength allocation policy used will have little effect on the blocking probability. It is at these high blocking probability values that having converters at all nodes (circuit-switched case) will help. However, it is unlikely that realistic networks will be designed to operate at this region.

In Figures 25 and 26 we compare the first-fit policy to the random policy with 4, 8, and 12 converters employed in the torus network (note that these values correspond to 16%, 32% and 48%, respectively, of the network nodes having converters). As we can see, using the first-fit policy is roughly equivalent to employing a significant number of converters in the network.

3.2 The NSFNET Topology

We have also considered a realistic example of a backbone network, namely, the NSFNET irregular topology shown in Figure 27. Since we will be using traffic data reported in [5], following that study, we have also augmented the 14-node NSFNET topology by adding two fictitious nodes, nodes 1 and 16 in Figure 27, to capture the effect of NSFNET's connections to Canada's communication network, CA*net. The resulting topology consists of 16 nodes and a total of 240 source-destination pairs. As in the previous subsection, we have decided to present detailed results for the call blocking probabilities of only a small number of pairs, and to summarize the results for the whole network. Specifically, we present detailed results for the blocking probabilities of calls involving nodes along the path (3,5,6,7,9,12,15,16). (We note, however, that the shortest path used by some of these calls is not a sub-path of (3,5,6,7,9,12,15,16); for instance, the shortest path for calls between nodes 3 and 15 is (3,5,11,15).) The 28 source-destination pairs in this path, along with the corresponding shortest path lengths and the labels used

We have used two different traffic patterns with the NSFNET topology. The first traffic pattern is similar to that used with the torus network. Specifically, the arrival rate λ_{sd} for a source-destination pair (s, d) is

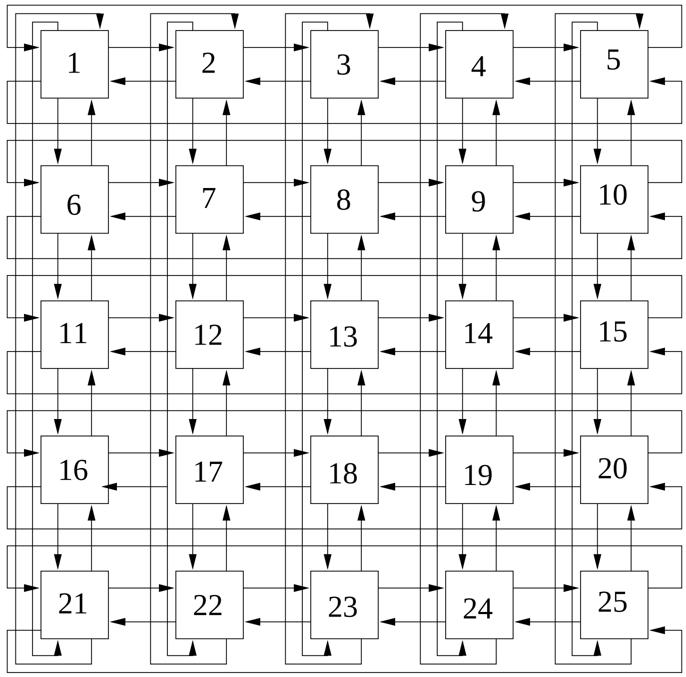

Figure 21: The 5×5 bidirectional mesh torus network

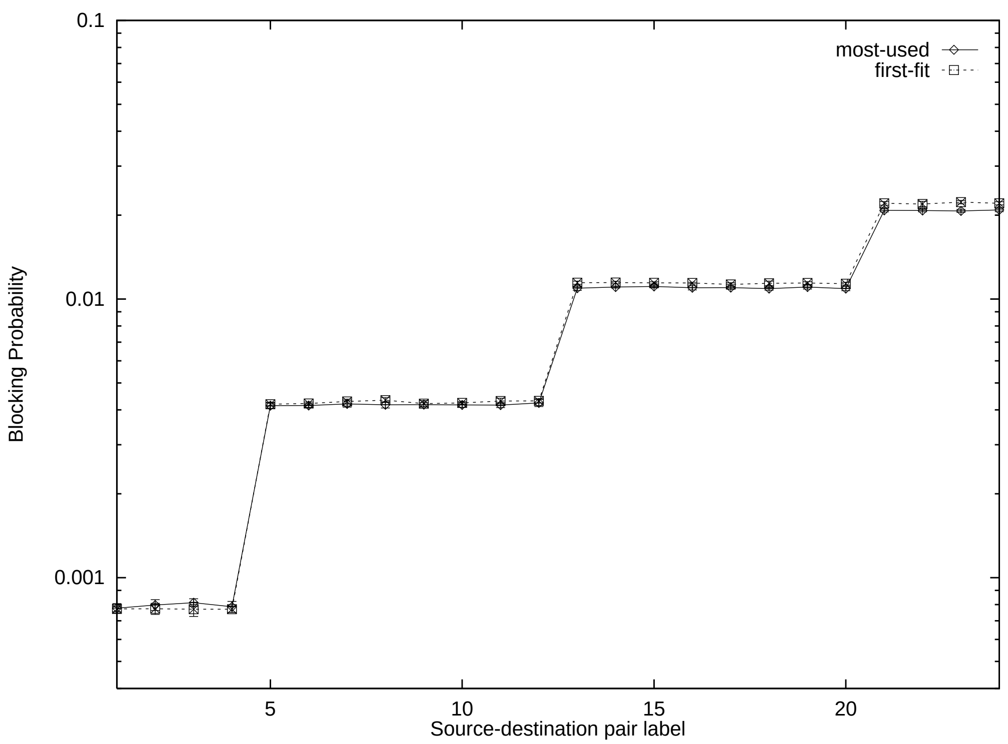

Figure 22: Most-used vs. first-fit allocation, 5×5 torus network, traffic pattern based on locality

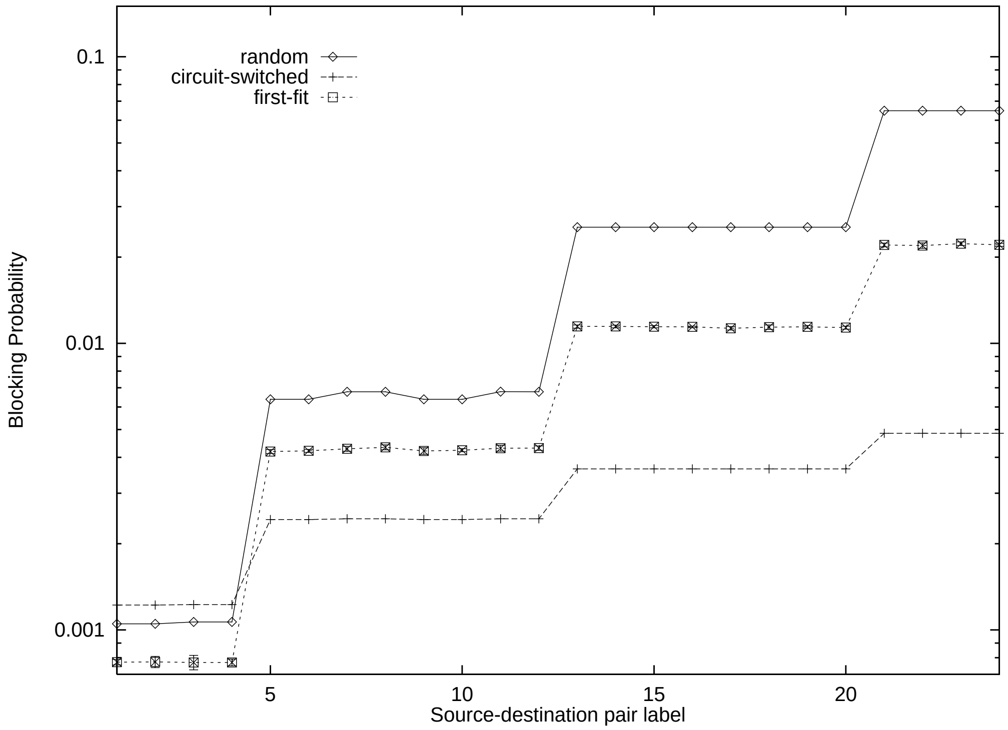

Figure 23: Policy comparison, 5×5 torus network, traffic pattern based on locality

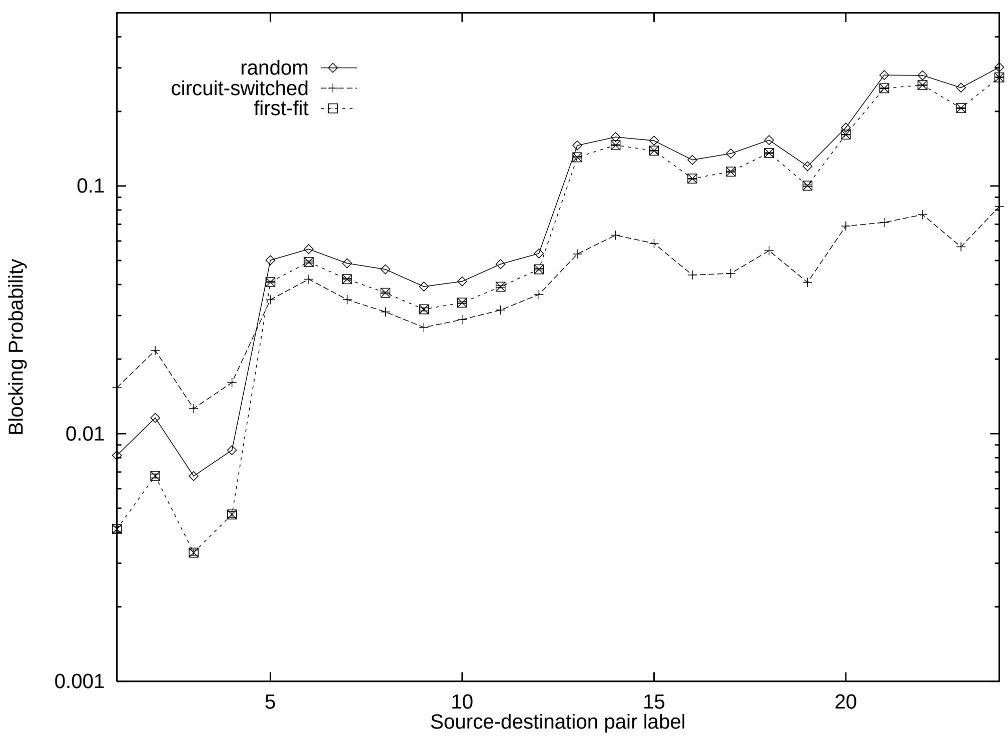

Figure 24: Policy comparison, 5×5 torus network, random traffic pattern

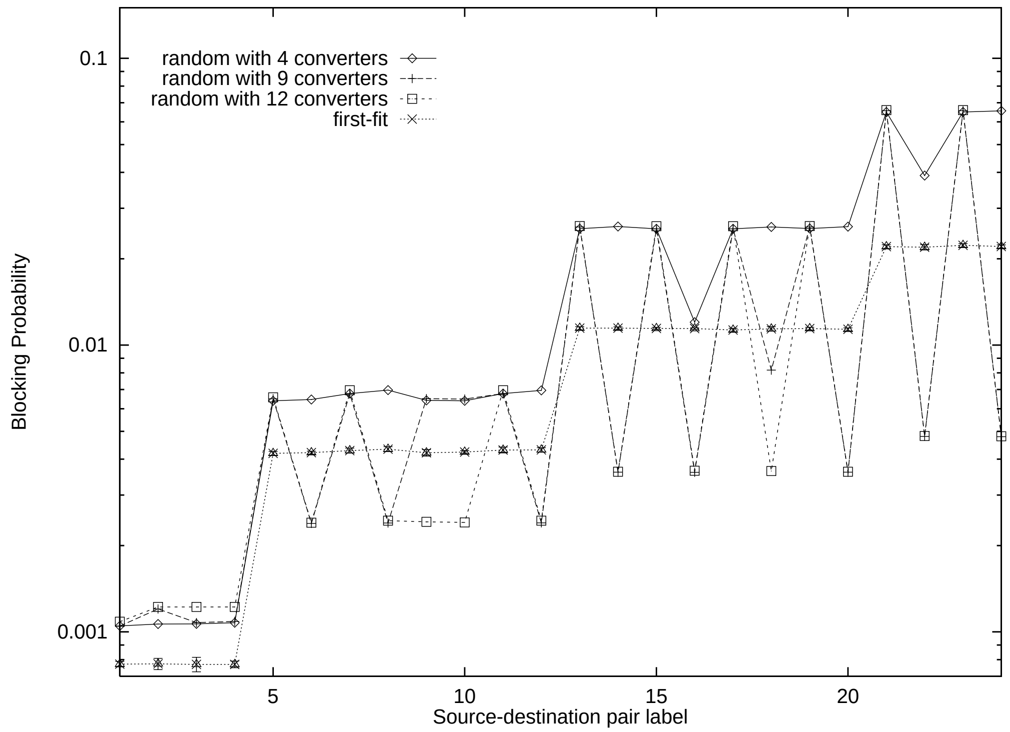

Figure 25: First-fit policy vs. random policy with converters, 5×5 torus network, pattern based on locality

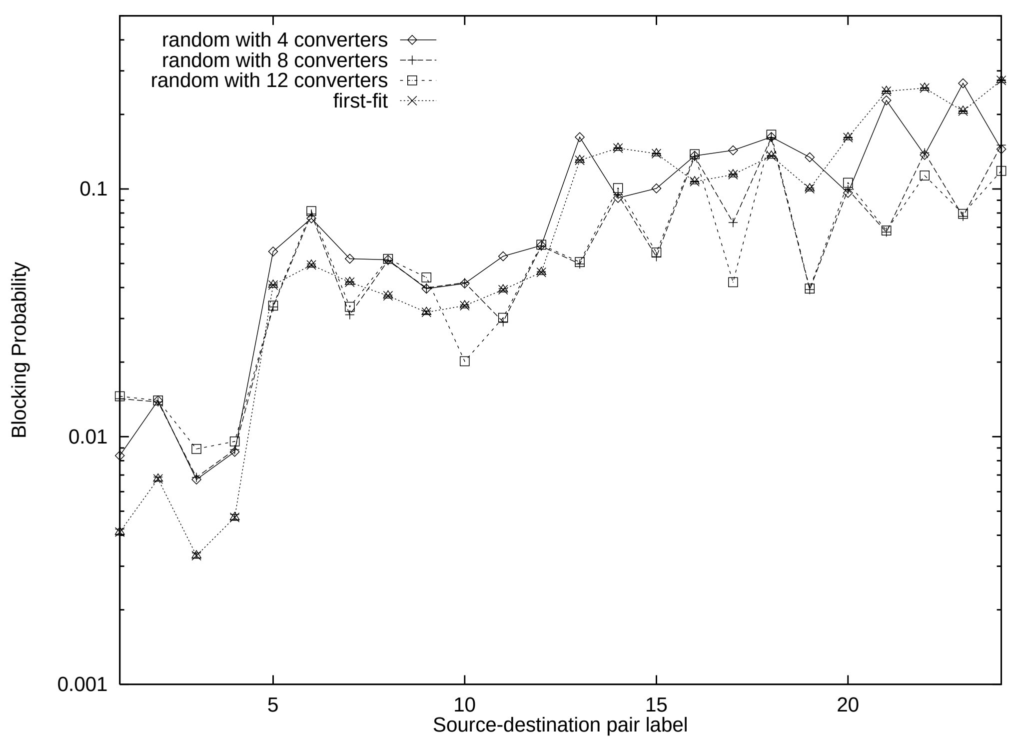

Figure 26: First-fit policy vs. random policy with converters, 5×5 torus network, random traffic pattern

given by (note that no shortest path is longer than 4 hops):

$$\lambda_{sd} = \begin{cases} 0.5, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 1 \\ 0.4, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 2 \\ 0.3, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 3 \\ 0.2, & \text{if the length of the shortest path from } s \text{ to } d \text{ is } 4 \end{cases} \tag{4}$$

The second traffic pattern was designed to reflect actual traffic statistics collected on the NSFNET backbone network, as reported in the traffic matrix in [5, Figure 6]. The data in this traffic matrix represent the measured number of bytes transferred from a node s to a node d in the NSFNET backbone within a certain 15-minute interval. Clearly, this data, collected over a packet-switched network, cannot be directly applied to a circuit-switched wavelength routing network, such as the one considered in this work. However, our intention is simply to capture the relative traffic demands among the different source-destination pairs. To this end, we first divide the entries of the matrix in [5, Figure 6] by the link capacity (T3 links) to obtain the "offered load" ρ_{sd} per source-destination pair. Since the resulting values are too small, we multiply them by a constant to obtain reasonable values for the offered load. Then, assuming that all calls have a mean holding time $1/\mu = 1$, the offered load values become the arrival rates λ_{sd} used in' the experiments. As a result, the relative values of these arrival rates reflect the relative traffic requirements among the different source-destination pairs according to the specific traffic pattern reported in [5].

Our results are presented in Figures 28-32. Figure 28 compares the first-fit to the most-used policies. Figures 29 and 30 demonstrate that the random and circuit-switched cases provide upper and lower bounds on the performance of the first-fit policy, while Figures 31 and 32 attempt to quantify the effect of the first-fit policy in terms of number of converters. The overall behavior of the graphs shown in these figures is very similar to that discussed earlier for the torus network and the single path cases, indicating that our observations and conclusions are valid for a wide range of network topologies and traffic patterns.

4 Concluding Remarks

We have studied the blocking performance of various wavelength allocation policies for various single path and network topologies and under various traffic patterns. Our conclusions can be summarized as follows:

- We have shown that the most-used and first-fit policies have very similar performance for all calls in a network, regardless of the number of hops used by the calls. Specifically, the two policies tend to favor calls using multiple paths at the expense of calls using a single path. This is a desirable feature, since calls traversing multiple paths experience the highest blocking probability. However, the most-used policy requires that the network nodes exchange information about the network-wide usage of wavelengths, while the first-fit policy only relies on a fixed ordering of wavelengths, making it significantly easier to implement.
- We have also demonstrated that the random policy and the circuit-switched case, for which analytical (exact or approximate) solutions exist for systems of large size, provide lower and upper bounds on the

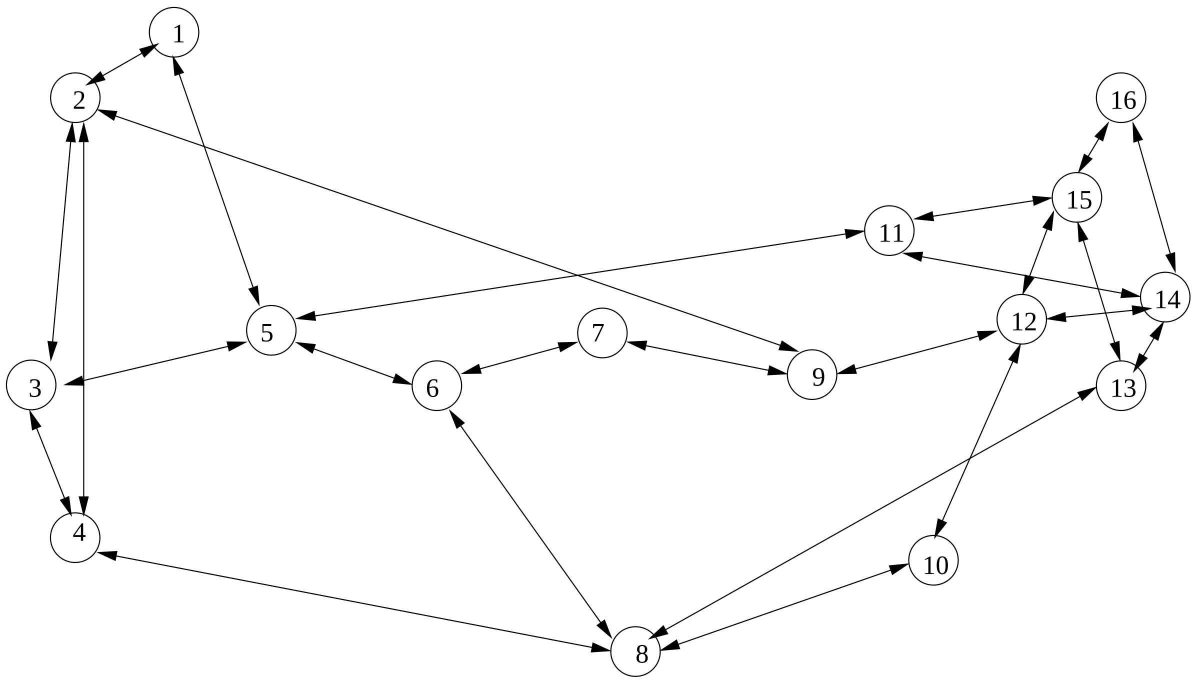

Figure 27: The NSFNET topology

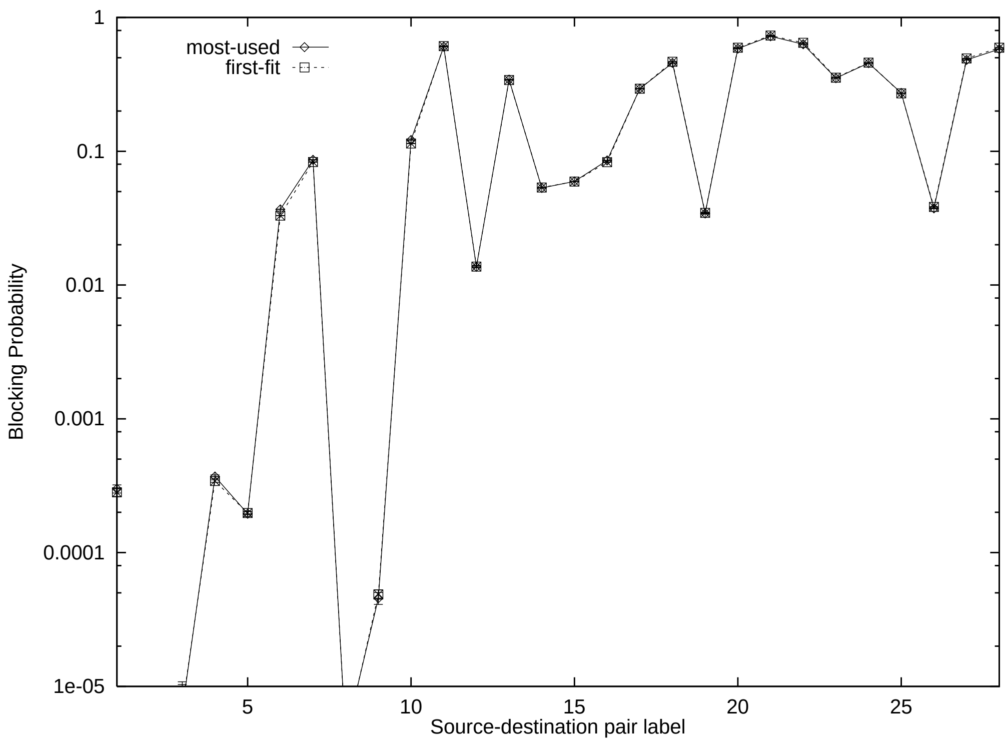

Figure 28: Most-used vs. first-fit allocation, NSFNET, pattern based on actual traffic

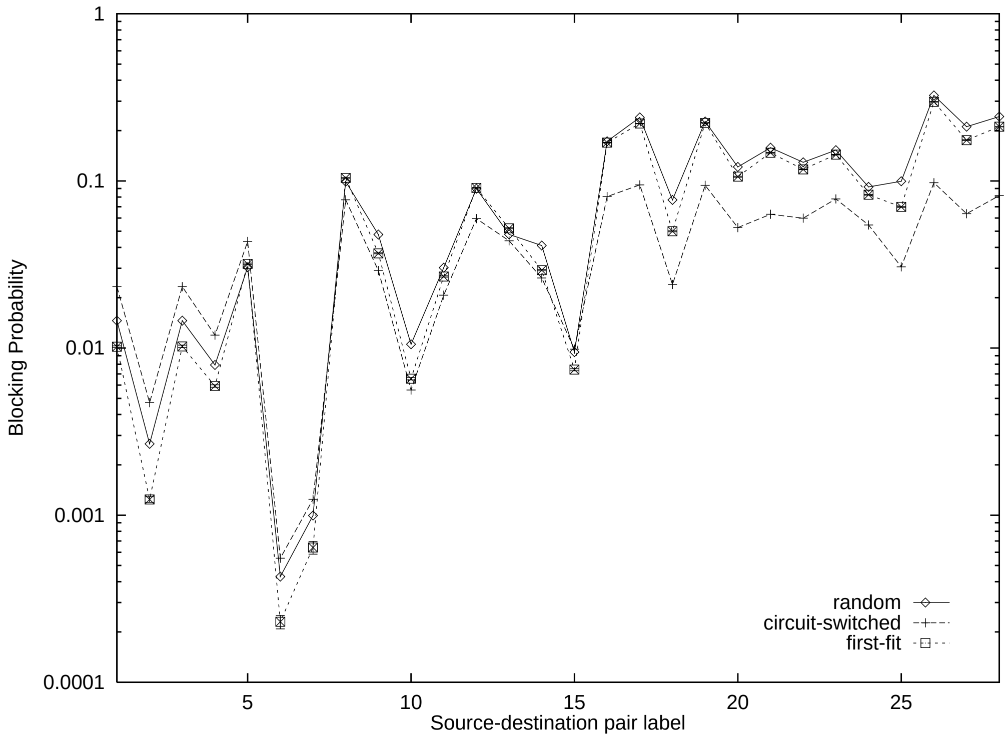

Figure 29: Policy comparison, NSFNET, traffic pattern based on locality

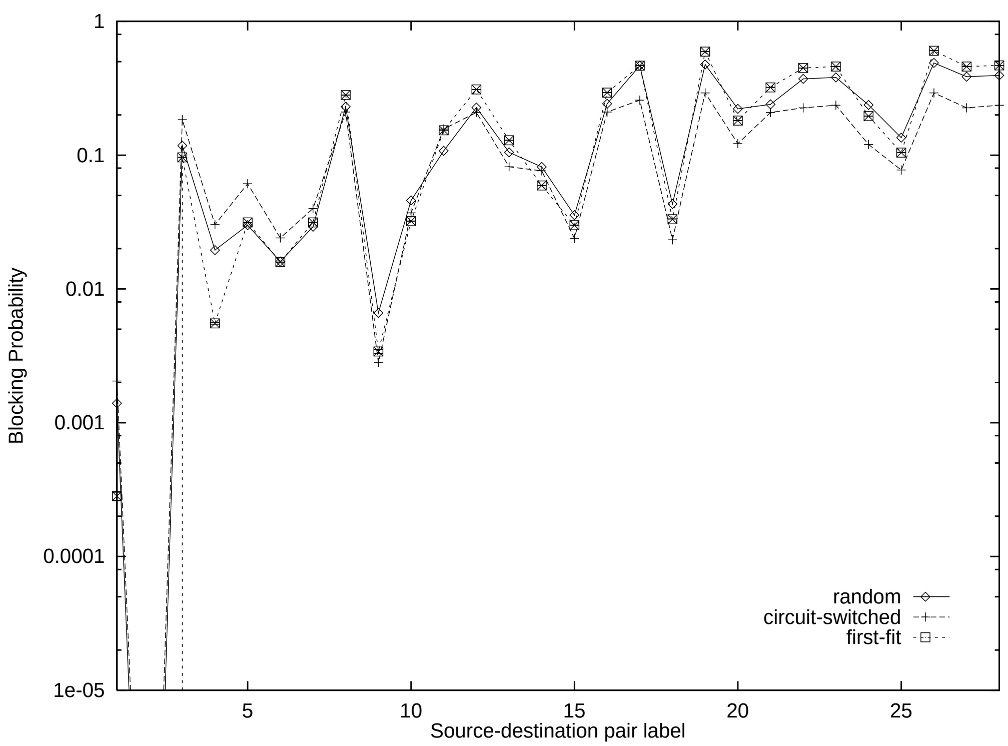

Figure 30: Policy comparison, NSFNET, pattern based on actual traffic

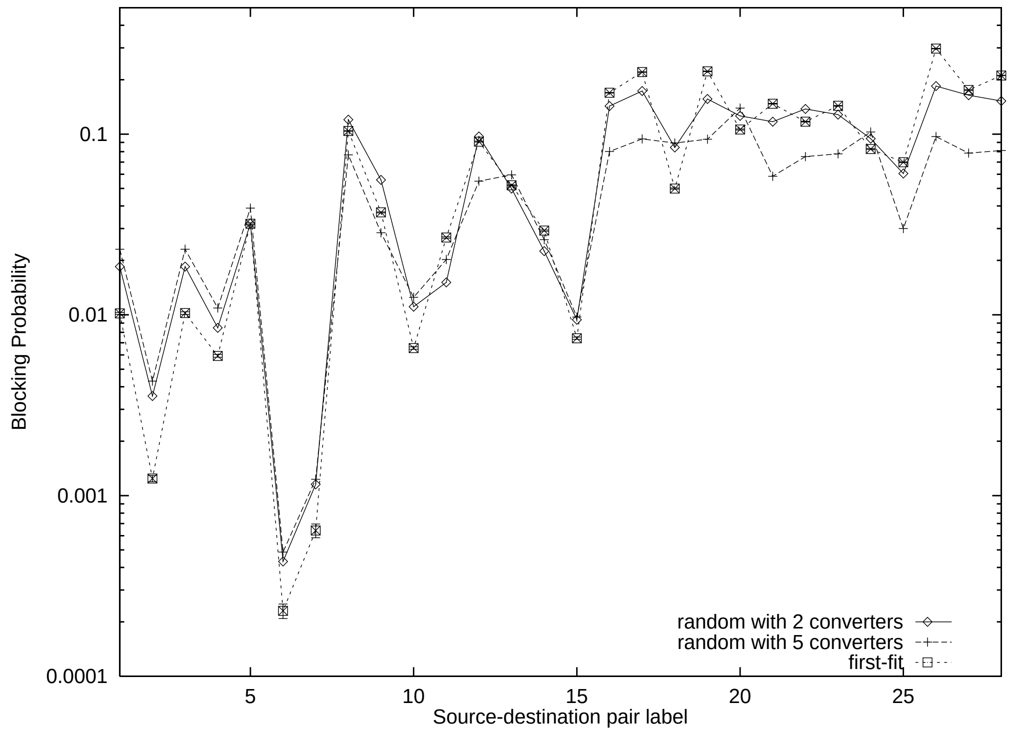

Figure 31: First-fit policy vs. random policy with converters, NSFNET, traffic pattern based on locality

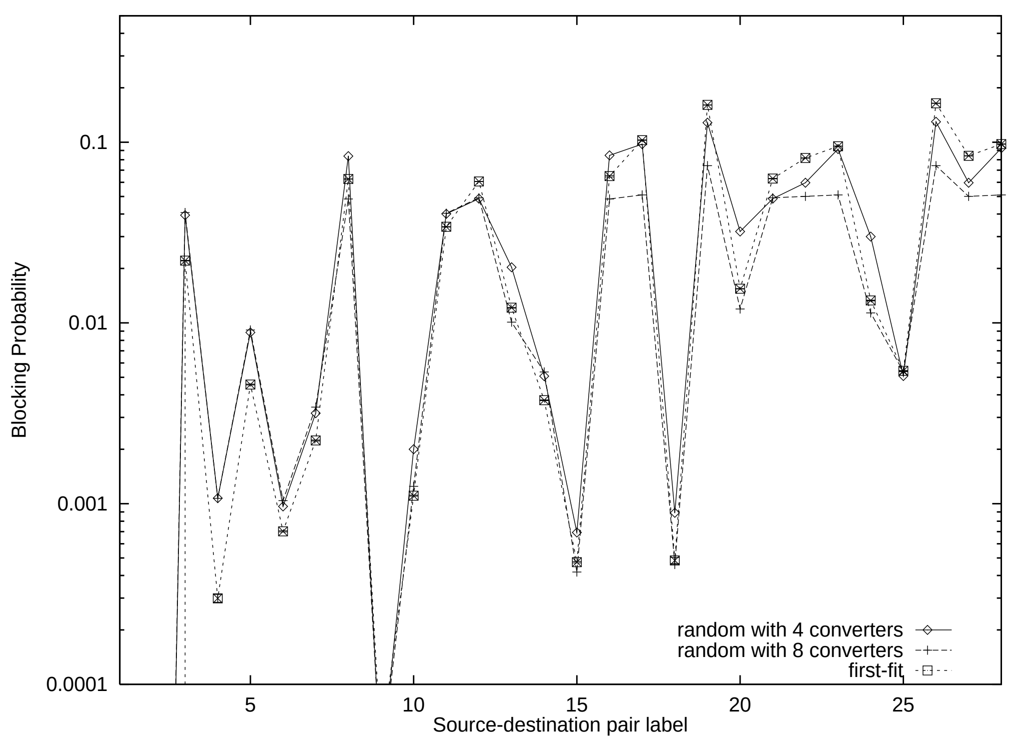

Figure 32: First-fit policy vs. random policy with converters, NSFNET, random traffic pattern

call blocking probability under the first-fit (or most-used) policy. Specifically, for calls using a single hop, the random policy provides a lower bound and the circuit-switched case provides an upper bound, while for calls using more than one hops the bounds are reversed.

- We have presented results which indicate that the effect of using the first-fit policy is "equivalent" to using the random policy but employing a number of converters (between 15% to 50% of the number of nodes) in the network. More importantly, in most cases, introducing the first-fit policy results in a decrease in the blocking probability of calls traveling over multiple hops to a level very close to the blocking probability experienced under the circuit-switched case. Note that, in terms of implementation, there is no significant difference between the first-fit and random policies. Consequently, the gains obtained by employing specialized (and expensive) hardware can be realized by making more intelligent choices in software.

- It also appears that the benefits of the first-fit policy diminish at high loads (blocking probability values of 0.1 or more). It is a these situations that employing converters would benefit calls traversing a large number of hops. However, the number of converters to be employed in this case must be very large, close to the number of nodes in the network, and even if all nodes contain converters the blocking probability will remain at (reduced but) high levels. Since it is unlikely that future wavelength routing networks will be designed to operate at such high call blocking probability, this result may not be of practical importance.

Overall, our results appear to contradict previous studies which have indicated that "sparse" wavelength conversion capabilities (i.e., selective placement of converters at a subset of network nodes) will be beneficial to wavelength routing networks. Those studies measured the improvement obtained by employing converters over the random wavelength allocation policy only, while we have shown that an equivalent improvement can be achieved merely by using appropriate allocation policies such as first-fit or most-used. While further investigation of the value of converters is warranted, our results can have a major impact on the direction of research and development in the area of wavelength converters.

References

[1] R. A. Barry and P. A. Humblet. Models of blocking probability in all-optical networks with and without wavelength changers. *IEEE Journal Selected Areas in Communications*, 14(5):858–867, June 1996.

[2] A. Birman. Computing approximate blocking probabilities for a class of all-optical networks. *IEEE Journal Selected Areas in Communications*, 14(5):852–857, June 1996.

[3] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: An approach to high bandwidth optical WANS. *IEEE Transactions on Communications*, 40(7):1171–1182, July 1992.

[4] R. Cruz, G. Hill, A. Kellner, R. Ramaswami, G. Sasaki, and Y. Yamabayashi (Eds.). Special issue on optical networks. *IEEE Journal Selected Areas in Communications*, 14(5), June 1996.

[5] B. Mukherjee *et al.* Some principles for designing a wide-area WDM optical network. *IEEE/ACM Transactions on Networking*, 4(5):684–696, October 1996.

[6] I. P. Kaminow *et al.* A wideband all-optical WDM network. *IEEE Journal Selected Areas in Communications*, 14(5):780–799, June 1996.

[7] R. E. Wagner *et al.* MONET: Multiwavelength optical networking. *Journal of Lightwave Technology*, 14(6):1349–1355, June 1996.

[8] S. P. Monacos *et al.* All-optical WDM packet networks. *Journal of Lightwave Technology*, 14(6):1356–1370, June 1996.

[9] M. Fujiwara, M. Goodman, M. O'Mahony, O. Tonguz, and A. Willner (Eds.). Special issue on multiple wavelength technologies and networks. *Journal of Lightwave Technology*, 14(6), June 1996.

[10] A. Girard. *Routing and Dimensioning in Circuit-Switched Networks.* Addison Wesley, Reading, MA, 1990.

[11] H. Harai, M. Murata, and H. Miyahara. Performance of alternate routing methods in all-optical switching networks. In *Proceedings of INFOCOM '97*, pages 517–525. IEEE, April 1997.

[12] E. Karasan and E. Ayanoglu. Effects of wavelength routing and selection algorithms on wavelength conversion gain in wdm optical networks. *IEEE/ACM Transactions on Networking*, 6(2):186–196, April 1998.

[13] M. Kovacevic and A. Acampora. Benefits of wavelength translation in all-optical clear-channel networks. *IEEE Journal Selected Areas in Communications*, 14(5):868–880, June 1996.

[14] A. Mokhtar and M. Azizoglu. Adaptive wavelength routing in all-optical netowrks. *IEEE/ACM Transactions on Networking*, 6(2):197–206, April 1998.

[15] B. Ramamurty and B. Mukherjee. Wavelength conversion in WDM networking. *IEEE Journal Selected Areas in Communications*, 16(7):1061–1073, Septemer 1998.

[16] S. Subramaniam, M. Azizoglu, and A. Somani. All-optical networks with sparse wavelength conversion. *IEEE/ACM Transactions on Networking*, 4(4):544–557, August 1996.

[17] S. Subramaniam, M. Azizoglu, and A. K. Somani. On the optimal placement of wavelength converters in wavelength-routed networks. In *Proceedings of INFOCOM '98*, pages 902–909. IEEE, April 1998.

[18] S. Subramanian, A. K. Somani, M. Azizoglu, and R. A. Barry. A performance model for wavelength conversion with non-poisson traffic. In *Proceedings of INFOCOM '97*, pages 500–507. IEEE, April 1997.

[19] Y. Zhu, G. N. Rouskas, and H. G. Perros. Blocking in wavelength routing networks, Part II: Mesh topologies. Technical Report TR-98-10, North Carolina State University, Raleigh, NC, October 1998.

[20] Y. Zhu, G. N. Rouskas, and H. G. Perros. Blocking in wavelength routing networks, Part I: The single path case. In *Proceedings of INFOCOM '99*. IEEE, March 1999. (To appear).