An Analysis of Oblivious and Adaptive Routing in Optical Networks With Wavelength Translation

Jonathan P. Lang, *Member, IEEE*, Vishal Sharma, *Member, IEEE*, and Emmanouel A. Varvarigos, *Member, IEEE*

***Abstract*—We present an analysis for both oblivious and adaptive routing in regular, all-optical networks with wavelength translation. Our approach is simple, computationally inexpensive, accurate for both low and high network loads, and the first to analyze adaptive routing with wavelength translation in wavelength division multiplexed (WDM) networks while also providing a simpler formulation of oblivious routing with wavelength translation. Unlike some previous analyses which use the link independence blocking assumption and the *call dropping* (loss) model (where blocked calls are cleared), we account for the dependence between the acquisition of wavelengths on successive links of a session's path and use a lossless model (where blocked calls are retried at a later time). We show that the throughput per wavelength increases superlinearly (as expected) as we increase the number of wavelengths per link, due both to additional capacity and more efficient use of this capacity; however, the extent of this superlinear increase in throughput saturates rather quickly to a linear increase. We also examine the effect that adaptive routing can have on performance. The analytical methodology that we develop can be applied to any vertex and edge symmetric topology, and with modifications, to any vertex symmetric (but not necessarily edge symmetric) topology. We find that, for the topologies we examine, providing at most one alternate link at every hop gives a per-wavelength throughput that is close to that achieved by oblivious routing with twice the number of wavelengths per link. This suggests some interesting possibilities for network provisioning in an all-optical network. We verify the accuracy of our analysis for both oblivious and adaptive routing via simulations for the torus and hypercube networks.**

***Index Terms*—Adaptive routing, all-optical networks, hypercube, multi-fiber networks, oblivious routing, performance analysis, torus, wavelength division multiplexing, wavelength translation.**

I. Introduction

THE TECHNOLOGY exists today to transmit gigabits of data per second over thousands of kilometers with extremely small loss. This has spurred a number of applications that were either infeasible or not cost-effective in the pre-gigabit era. The rapid demand of these emerging gigabit-per-user applications, however, has outstripped the gains possible via the traditional approach of building faster time division multiplexed (TDM) networks. This has increased interest in building *all-optical networks* where wavelengths, rather than timeslots, are switched.

In an all-optical wavelength division multiplexed (WDM) network, connection establishment for a session involves two phases: the selection of a *route*, or sequence of hops that the session must traverse, and for each hop along the route, the selection of a wavelength on which the session will be carried for that hop. The *path* of a session is the sequence of link-wavelength pairs traversed by it. Path selection, therefore, involves *routing* and *wavelength assignment*, both of which may be either *oblivious* or *adaptive*. In *oblivious* (or *static*) routing, the route is selected at the source and is independent of the state (loading or congestion) of the network, while in *adaptive* (or *dynamic*) routing, the route is selected either at the source or on a hop-by-hop basis, based on the state of the network.

A critical functionality for the improved performance of multihop WDM networks is *wavelength translation* [1], which is the ability of network nodes to switch data from a wavelength λ_i on an incoming link (the *incoming wavelength*) to a wavelength $\lambda_j,\ j \neq i$ on an outgoing link (the *outgoing wavelength*). Three natural classes of wavelength-routing nodes in this context are: 1) nodes with *full-wavelength translation* capability (see, for example, [2]–[4]), which can switch an incoming wavelength to any outgoing wavelength; 2) nodes with *limited-wavelength translation* capability (see, for example, [5]–[9]), which can switch an incoming wavelength to a subset of the outgoing wavelengths; and 3) nodes with *no-wavelength translation* capability (see, for example, [6], [10]–[14]), which can switch each incoming wavelength only to the same outgoing wavelength, the so-called wavelength-continuity constraint. The requirement of wavelength continuity restricts the routing flexibility and increases the probability of call blocking [2]. In this paper, we assume that nodes have full-wavelength translation capability. We examine, however, the tradeoffs involved when a switch with $k \times k$ wavelength translation capability is replaced by $k/n,\ n < k$ simpler switches, each having $n \times n$ wavelength translation capability.

A. Previous Work

Routing and wavelength assignment in WDM networks has recently received considerable attention. The first body of work in this area has concentrated on the performance of full- or no-wavelength translation with oblivious routing. Although this initial work correctly identifies several parameters that affect the performance of wavelength translation (such as path length,

Manuscript received May 12, 1999; revised December 5, 1999; approved by IEEE/ACM Transactions on Networking Editor G. Sasaki. This work was supported by the Defence Advanced Research Projects Agency (DARPA) under the MOST project.

J. P. Lang is with Calient Networks, Goleta, CA 93117 USA (e-mail: jplang@calient.net).

V. Sharma is with Metanoia, Inc., San Jose, CA 95134 USA (e-mail: v.sharma@ieee.org).

E. A. Varvarigos was with Delft University of Technology, Delft, The Netherlands and the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA. He is now with the Department of Computer Engineering and Informatics, University of Patras, Patras, Greece (e-mail: manos@ceid.upatras.gr).

Publisher Item Identifier S 1063-6692(01)06856-X.

number of wavelengths, switch size, network topology, and interference length), and provides useful qualitative insights into network behavior, several difficulties remain. One problem is accurately accounting for the load correlation between the wavelengths on successive links of a session's path. Kovačević and Acampora [3] provided a model to compute the approximate blocking probability in WDM networks with and without wavelength translation. As they point out, however, their model is inappropriate for sparse networks because it uses the link independence blocking assumption, which does not consider the dependence between the acquisition of wavelengths on successive links of a session's path. Barry and Humblet [2] presented an analysis that takes the link-load dependence partially into account, but they assumed that a wavelength is used on a link independently of other wavelengths. Their simplified model makes good qualitative predictions of network behavior (predicting even some nonobvious behavior observed in simulations [15]), but is unable to predict the behavior of simulations with numerical accuracy. The analysis by Birman [12], which assumes networks without wavelength translation, uses a Markov chain model with state-dependent arrival rates and is more accurate than the previous models, but involves modified reduced-load approximations and is computationally intensive for networks with more than three hops per route and a modest number of wavelengths per link. All the aforementioned analyses use the *call dropping* (loss) model, where blocked calls are cleared, which tends to overestimate the achievable throughput for a given blocking probability by favoring short connections.

The analysis presented in [8] for limited wavelength translation accounts partially for the link-load dependence, and maintains fairness to all connections by retrying blocked sessions at a later time (lossless model). This analysis can be applied to full-wavelength translation, but the number of states grows exponentially with the degree of translation, and is impractical when the number of wavelengths per link is large. Recent work by Zhu *et al.* [16] develops an exact Markov process model to obtain the call-blocking probability for an h-hop path that takes link-load correlation and nonuniform traffic into account, but requires h^2 state variables and is impractical for large networks.

A second body of work deals with the performance evaluation of routing and wavelength assignment algorithms in all-optical networks. Karasan and Ayanoglu [17] analyzed the first-fit wavelength assignment strategy in a network with no-wavelength translation and fixed shortest-hop routing. They also proposed an adaptive routing and wavelength assignment (RWA) algorithm and evaluated its performance via simulations. Mokhtar and Azizoglu [18] also proposed and simulated several adaptive RWA algorithms for networks with no-wavelength translation, and analyzed oblivious alternate routing, where several disjoint paths are tried, using a fixed-order wavelength search. Harai *et al.* [19] analyzed oblivious alternate routing with fixed wavelength assignment and no-wavelength translation, and recently [9], they analyze oblivious alternate routing with various wavelength assignment schemes for networks with limited wavelength translation. Although these previous works provide valuable insight into the behavior of adaptive routing in all-optical networks, they continue to be based on simulations. The analyses given are only for oblivious (fixed or alternate) routing with fixed wavelength assignment schemes, and are still fairly complex, requiring the solution of a series of nonlinear Erlang maps and the predefining and ordering of all alternate paths between each source–destination pair, which becomes impractical for large networks with several wavelengths per link. Furthermore, the algorithms proposed require information on global wavelength utilization, assuming either a periodic exchange of such information [18], or a *centralized* network controller [17], [19].

B. Model

The analysis that we present for studying oblivious and adaptive routing in regular, all-optical networks with wavelength translation assumes a *distributed* network model. The routing decision is made locally at each node, using information only about the state of its own outgoing links and wavelengths. We do not require that the alternate paths between a source–destination pair be link disjoint [18], [19], instead allowing links (and wavelengths) to overlap between alternate paths.

We first present a general analysis applicable to any regular topology that employs either oblivious or adaptive routing. Our analysis holds for any vertex and edge-symmetric topology, and with modifications, to any vertex symmetric (but not edge-symmetric) topology. The analysis that we develop is simple, computationally inexpensive, and accurate for both low and high network loads, overcoming many of the difficulties of the first body of work highlighted earlier. We then apply our analysis to study the performance of oblivious and adaptive routing with wavelength translation in the torus and hypercube topologies. In our model, new sessions with uniformly distributed destinations arrive independently at each node of the network according to a Poisson process. A circuit is established by sending a setup packet along a shortest route from the source to the destination, and its success is assumed to be instantaneously known at the source. At each hop, the setup packet randomly selects a wavelength from among the available wavelengths, and if it is successful in establishing a connection, the wavelengths required by the session are reserved for the session duration; otherwise, the session is randomly assigned a new time at which to try. In oblivious routing, the route is selected at the source, while in adaptive routing, the intermediate links (and wavelengths) of the path are determined dynamically on a hop-by-hop basis, depending on link utilization; we require that, at each node, an outgoing link be selected from among the subset of outgoing links that lie on a shortest route to the destination. In a large mesh, most intermediate nodes have two outgoing links lying on a shortest route. In a hypercube, there are i outgoing links lying on a shortest route when the packet is at a distance i from its destination.

The capacity of each link is divided into k wavelengths, and each node has full-wavelength translation capability. We model an outgoing link of a node with k wavelengths per link by an auxiliary $M/M/k/k$ queuing system. Using the occupancy distribution of this system, we derive a closed-form expression for the probability $P_{\rm succ}$ of successfully establishing a circuit. To evaluate this probability, we do not use the link independence blocking assumption, but instead account partially for the dependence between the acquisition of successive wavelengths on the path followed by a session. Our analysis is general, computationally inexpensive (it avoids, for example, Erlang fixed-point or reduced load approximations, which, although they are asymptotically exact [20] in the regime of infinite link capacity or number of wavelengths and have good performance in the finite case [21], entail a high computational complexity), and scales easily for larger network sizes and arbitrary k. Our analysis applies also to both oblivious and dynamic routing, and for *regular networks*, provides accurate estimates of blocking and throughput at a computational cost smaller than in some of the approximations considered previously [22], [23]. Finally, we note that our analysis applies equally well to multifiber networks, with no-wavelength translation.

We examine how the extent of improvement in achievable throughput, for a fixed $P_{\rm succ}$, depends on the number of wavelengths k per link, and on the number of links l that may be tried at each hop (which we call the *routing flexibility*). This is important because it impacts the cost and complexity of the switch. Increasing the routing flexibility l increases the switch complexity and delay. Similarly, with full-wavelength translation, increasing the number of wavelengths k per link increases hardware complexity, and may be difficult to realize with current technology. We find that although the throughput per wavelength increases superlinearly with k, the incremental gain in throughput per wavelength (for a fixed $P_{\rm succ}$) saturates rather quickly to a linear increase. We also see that when the routing flexibility l is varied, the largest incremental gain in throughput per wavelength occurs when l is increased from one to two. We also compare the performance obtainable with a certain number of wavelengths k with that obtainable with a certain routing flexibility l. For the torus and hypercube topologies, we find that for a fixed $P_{\rm succ}$, a system with k wavelengths per link and only one alternate choice of an outgoing link (i.e., $l = 2$) gives a per-wavelength throughput that is close to that achieved by a system using oblivious routing with $2k$ wavelengths per link, with only a small additional improvement as l is increased further. This tradeoff between routing flexibility and degree of wavelength translation, which has a significant impact on network dimensioning, has not been studied before in the literature. A byproduct of this research is a better understanding of various oblivious routing schemes. We show, for example, that for the torus network, X–Y routing performs better than Zig-Zag routing. The above observations imply several interesting alternatives for the provisioning and expansion of all-optical networks, some of which we discuss in Section V.

The organization of the remainder of the paper is as follows. In Section II, we present a general analysis for any regular network with wavelength translation using either oblivious or adaptive routing. In Section III, we apply it to the torus network, where we examine two oblivious routing schemes (X–Y routing and Zig-Zag routing), and a shortest-hop adaptive routing scheme. In Section IV, we apply our analysis to the hypercube network for both oblivious and adaptive routing. In Section V, we present results for the probability of success obtained from our analysis and compare them to those obtained via simulations, and we discuss our results. In Section VI, we present our conclusions.

II. A General Methodology

In this section, we present a general methodology for analyzing oblivious and adaptive routing in regular networks with wavelength translation. We define a *regular network* to be one which is either vertex (but not edge) symmetric or is both vertex and edge symmetric. In Sections III and IV, we apply our methodology to analyze the performance of the torus and hypercube networks, respectively. Our choice of the torus and hypercube topologies reflects our interest in analyzing two popular topologies with very different characteristics. The torus is a sparse topology with a small (fixed) node degree and rather large diameter, while the hypercube is a dense topology, with node degree and diameter that increase logarithmically with the number of nodes. The general formulation that we develop provides a method to analyze other regular topologies that have been proposed for building all-optical networks, such as the family of banyan networks, e.g., shufflenet [24], [25], and wrapped butterfly networks [8].

In our model, external session requests are generated independently at each node of the network according to a Poisson process with rate λ sessions per unit time, and their destinations are uniformly distributed over all nodes, except for the source node. The holding time, or duration, of a session is exponentially distributed with unit mean ($1/\mu = 1$). Connections are established by transmitting a setup packet from the source to the destination. In a circuit-switched network, the duration of a session is typically much longer than the propagation and processing time of a setup packet along the route (otherwise circuit switching would be rather inefficient); the blocking of new session requests, therefore, is due primarily to the presence of existing sessions. Thus, we simplify modeling by assuming that new sessions are blocked only by pre-existing sessions, and that a source receives *instantaneous* feedback about whether or not a given session request can be satisfied. This allows us to concentrate on the main features of the routing schemes without having to focus on any specific implementation of the signaling/control mechanisms and without having to account for second-order effects such as the blocking of sessions due to resources consumed by partial reservations that were not finally completed due to limited capacity at downstream nodes. Indeed, earlier analyses [26] have shown that when the connection setup time is a small fraction of the session holding time (as would be case in circuit-switched networks), the effect of overhead due to reservations is rather small. By deploying intelligent, distributed reservation and connection control schemes, such as those proposed in [7], [27], and [28], the effect of these overheads can be further minimized.

In our scheme, therefore, the setup process works as follows. If the setup packet is successful in establishing the circuit, the wavelengths required by it are reserved for the duration of the session. Otherwise, the session is blocked and is assigned a new time at which to try. The duration of this interarrival time is drawn from a Poisson distribution with an arrival rate that is much lower than (say, $1/10$th) the arrival rate of the original

Poisson stream (correspondingly, its interarrival time is much larger than the interarrival time of the original Poisson stream). This is done to ensure that reinserting the blocked session into the input stream produces a combined process of exogenous arrivals and retrials that can be approximated as a Poisson process, and that all sessions are eventually served. By contrast, in the call-dropping model used in previous analyses, sessions with longer route lengths are dropped with a higher probability (unfairly treating such connections), and the maximum throughput is overstated, especially at higher loads. In the subsequent sections, we define an auxiliary system that we use to model an outgoing link at a node, and we obtain the probability of successfully establishing a circuit.

A. The Auxiliary System

We focus on setup packets emitted on an outgoing link L of a node s, and define the *type* τ of a setup packet according to whether it belongs to a session originating at the node or according to the incoming link upon which it arrives. For regular networks, it is useful to partition the set of incoming links at a node into groups in the following way. We denote the network $G = (N, A)$, where N is the set of nodes, and A is the set of links. A $1-1$ function T defined over N will be called an automorphism if for every link $v \in A$ we have that $T(v) \in A$. T will be called a *fixed* automorphism for link L if it maps L to itself. We say that two incoming links l_1 and l_2 of node s belong to the same *spatial group*, with respect to L if there exists a fixed automorphism T for L, such that $T(l_1) = l_2$. Intuitively, this means that if we focus on outgoing link L, then there is symmetry between l_1 and l_2. We use this mapping to partition the $r-1$ incoming links of a node (except for link L) into q, $1 \leq q \leq r-1$, different groups (or types), so that the links of each group q have the same spatial relationship with respect to outgoing link L. The total number of incoming links of type q is denoted by M_q [see also Fig. 1(a)]. Originating setup packets that are emitted on link L are defined as being of type $\tau = 0$, while transit setup packets are defined as being of type $\tau = 1, 2, \ldots, q$ if the incoming link over which they arrive is of type τ. We also let $\gamma(\tau)$ denote the rate per unit of time at which setup packets of type τ are emitted on an outgoing link.

We denote the state of an outgoing link by the vector $\overline{S} = (S_0, S_1, \ldots, S_q)$, where S_0 is the number of originating sessions on the link, and S_τ, $\tau = 1, 2, \ldots, q$, is the number of transit sessions of type τ using the link. The set of feasible states of the outgoing link is given by $\mathcal{F} = \{\overline{S}\colon |\overline{S}| \leq k\}$, where $|\overline{S}| = S_0 + S_1 + \cdots + S_q$ and k is the number of wavelengths per link. We let $\pi(\overline{S})$ be the steady-state probability that an outgoing link is in state $\overline{S}$. We approximate $\pi(\overline{S})$ as the stationary distribution of an auxiliary $M/M/k/k$ queuing system Q, defined as follows [see Fig. 1(b)]. Customers of type τ, $\tau = 0, 1, \ldots, q$ arrive to the system Q according to a Poisson process with rate $\gamma^*(\tau)$, and ask for a server from among the k identical servers. If all k servers are busy, the customer is dropped, never to appear again. We require that the rate at which customers of type τ are accepted in the auxiliary system Q be the same as the rate $\gamma(\tau)$ at which setup packets of type τ are emitted on an outgoing link in the actual system. [Note that $\gamma(\tau)$ depends on the particular network topology and the routing algorithm used; in

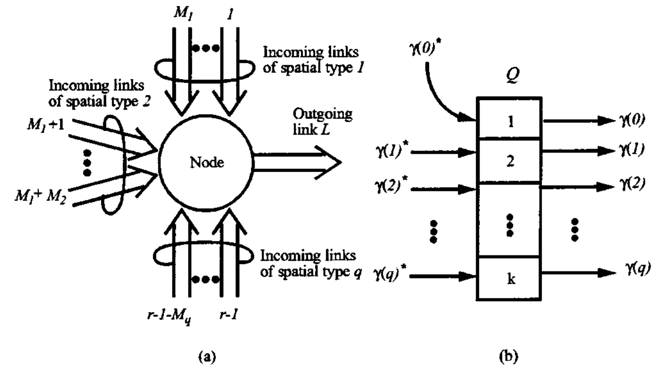


Fig. 1. (a) Illustration of how the $r-1$ incoming links (not including link L) are related to the outgoing link L at a node of the network, where each link has k wavelengths. For example, in the $p \times p$ torus network (see Section III) the two incoming links that lie along the dimension perpendicular to L will be one spatial type ($\tau = 1$) and the incoming link that lies along the same dimension as L will be another spatial type ($\tau = 2$). In the 2^r-node hypercube network (see Section IV), all of the $r-1$ incoming links are the same spatial type due to the symmetry of the hypercube network. (b) The auxiliary $M/M/k/k$ queuing system $\mathcal{Q}$.

Sections III and IV, we show how to calculate $\gamma(\tau)$ for the torus and hypercube networks.] For this to hold, we must have

$$\gamma^*(\tau) = \frac{\gamma(\tau)}{1 - \displaystyle\sum_{\overline{S}\colon |\overline{S}|=k} \pi(\overline{S})}. \tag{1}$$

To calculate the steady-state probabilities $\pi(\overline{S})$ for all feasible states, we write down the global balance equations for the Markov chain corresponding to the auxiliary system Q

$$\pi(\overline{S}) = \frac{\displaystyle\sum_{\tau} \left\{\gamma^*(\tau)\pi\left(\overline{S} - e_\tau\right) + \mu(S_\tau + 1)\pi\left(\overline{S} + e_\tau\right)\right\}}{\displaystyle\sum_{\tau} \left\{\mu S_\tau + \gamma^*(\tau) I_{|\overline{S}|}(k)\right\}} \tag{2}$$

where

$$I_a(b) = \begin{cases} 1, & \text{if } a \neq b \\ 0, & \text{otherwise.} \end{cases}$$

e_τ is a unit vector of dimension q whose τth component is one, and "+" (or "−") corresponds to componentwise addition (or subtraction). Equations (1) and (2), together with the normalization condition $\sum_{\overline{S} \in \mathcal{F}} \pi(\overline{S}) = 1$, can be solved iteratively to obtain the steady-state probabilities $\pi(\overline{S})$ and the rates $\gamma^*(\tau)$, $\tau = 0, 1, \ldots, q$.

B. The Success Probability

We now examine the probability of success P_{succ} for a new session and the probability of success $\overline{P}_{\mathrm{succ}}$ for a random trial (both new and reattempting sessions) for both oblivious and adaptive routing. To determine the success probabilities, we first find an approximate expression for the probability $P_{\mathrm{succ}}(s, d)$ that a new session with a given source–destination pair (s, d) successfully establishes a circuit, and then average over all source–destination pairs to determine the average probability that a session is successful.

1) Oblivious Routing: We first study oblivious (or nonadaptive) routing, where the route followed by a session is chosen at the source and is independent of the state of the links. In oblivious routing with full-wavelength translation, a session is blocked and scheduled to retry only if all k wavelengths on the desired outgoing link are unavailable, where we assume that a setup packet selects the outgoing wavelength from among the available wavelengths on the link with equal probability.

The path followed by a session with source destination pair $(s,\, d)$ consists of an originating node followed by a sequence of transit nodes. The probability α_0 that a wavelength on the outgoing link of the originating node is available is given by

$$\alpha_0 = \sum_{\overline{S}:\, |\overline{S}|<k} \pi(\overline{S}). \tag{3}$$

At each transit node, the probability that a wavelength is available on an outgoing link L given that a transit setup packet of type τ arrived on link $L-1$ can be found to be

$$\alpha_\tau = \frac{\displaystyle\sum_{\overline{S}:\, |\overline{S}|<k} \pi(\overline{S}) \left(1 - \frac{S_\tau}{kM_\tau}\right)}{\displaystyle 1 - \sum_{\overline{S}:\, |\overline{S}|\le k} \pi(\overline{S}) \frac{S_\tau}{kM_\tau}} \tag{4}$$

where M_τ is the number of input links of type τ. In other words, α_τ is the probability $|S| < k$ given that a wavelength was available on an incoming link. The numerator in (4) is the sum of all of the state probabilities where at least one wavelength on outgoing link L is available, conditioned on the fact that an incoming wavelength on link $L-1$ is available. The multiplicative factor $(1 - S_\tau/kM_\tau)$ is needed because the wavelengths in use on link L cannot be in use by sessions from the particular wavelength on link $L-1$ upon which the transit setup packet reserved the resources on the previous node. The denominator is one minus the sum of the probabilities of all states where link L is unavailable, conditioned on the fact that the incoming wavelength on link $L-1$ is available.

In writing (3) and (4), we do *not* assume that the probabilities of acquiring wavelengths on successive links of a session's path are independent. Instead, we account partially for the dependence between the acquisition of successive wavelengths on a session's path by using the approximation that the probability of acquiring a wavelength on link L depends on the availability of a wavelength on link $L-1$ (in reality this probability depends, very weakly, on the availability of a wavelength on every link $1, 2, \ldots, L-1$ preceding link L). The simulation results presented in Section V demonstrate that, while the link independence blocking approximation used in other analyses can lead to a very poor prediction of the success probability, our approximation is a very good one.

The (conditional) probability of successfully establishing a circuit is then given by

$$P_{\text{succ}}(s,\, d) = \alpha_0 \prod_{\tau=1}^{q} \alpha_\tau^{h_\tau(s,\, d)}$$

where $h_\tau(s,\, d)$ is the number of hops on which the transit session is of type τ for a particular source–destination pair $(s,\, d)$, and the product is taken over all types $\tau, \tau = 1, \ldots, q, q \le r-1$, of transit sessions. For uniformly distributed destinations, the average probability of success P_{succ} for a new arrival can be written as

$$P_{\text{succ}} = \frac{1}{N(N-1)} \sum_{(s,\, d)} P_{\text{succ}}(s,\, d) \tag{5}$$

where N is the total number of nodes in the network.

In our model, sessions that are not successful in establishing a circuit are blocked and reinserted into the input stream. Since sessions with longer routes are blocked and reattempted with higher probability than sessions with shorter routes, the destination distribution of the connection attempts (both new and reattempting sessions) may no longer be the same as that of only new arrivals. Indeed, the number of sessions with source–destination pair $(s,\, d)$ will be inversely proportional to the success probability $P_{\text{succ}}(s,\, d)$. Thus, when calculating the success probability of a random connection attempt (averaged over all trials, new and reattempting), the success probability $P_{\text{succ}}(s,\, d)$ must be weighted by the fraction $w(s,\, d)$ of sessions in the total mix that wish to go from s to d. Hence, the success probability $\overline{P}_{\text{succ}}$ of a random connection attempt (averaged over all trials, both new and reattempting) can be written as

$$\overline{P}_{\text{succ}} = \sum_{(s,\, d)} P_{\text{succ}}(s,\, d) w(s,\, d)$$

where

$$w(s,\, d) = \frac{P_{\text{succ}}^{-1}(s,\, d)}{\displaystyle\sum_{(s,\, d)} P_{\text{succ}}^{-1}(s,\, d)}$$

is a weighting factor that accounts for the changed distribution of sessions in the overall mix due to the retrials. This reduces to

$$\overline{P}_{\text{succ}} = \frac{N(N-1)}{\displaystyle\sum_{(s,\, d)} P_{\text{succ}}^{-1}(s,\, d)}. \tag{6}$$

Note that $\overline{P}_{\text{succ}}$ is the harmonic mean of the $P_{\text{succ}}(s,\, d)$ over all pairs $(s,\, d)$, $s \ne d$, while P_{succ} is the arithmetic mean.

2) Adaptive Routing: We now consider adaptive routing, where at each hop, a link is selected at random from among all the outgoing links that lie on a shortest route to the destination. If all of the k wavelengths on the chosen link are unavailable, an alternate link lying on the shortest route to the destination is tried. This process continues until either an available link is found, or all of the alternate links have been examined. The number of alternate links that lie on a shortest route to the destination may change as the setup packet progresses toward its destination. Furthermore, a limit on the number of alternate links that are examined could be used to reduce further congestion, or the processing overhead at an intermediate node.

We let l be the number of outgoing links that a session may try at each hop, which we refer to as the routing flexibility. The number of *feasible* outgoing links at a node t is given by

$\min(l, n_{t,d})$, where $n_{t,d}$ is the number of outgoing links at node t that lie on a shortest route to the destination d. Therefore, a session currently at node t will be blocked and scheduled to retry only if all of the k wavelengths on each of its feasible outgoing links are unavailable.

In adaptive routing, we define only two types of setup packets: originating type $\tau = 0$, and transit type $\tau = 1$. The probability α_0 that at least one wavelength on $\min(l, n_{s,d})$ outgoing links at the origin is available is given by

$$\alpha_0(\min(l, n_{s,d})) = 1 - \left(1 - \sum_{\overline{S}:\, |\overline{S}|<k} \pi(\overline{S})\right)^{\min(l, n_{s,d})} \tag{7}$$

where $n_{s,d}$ is the number of outgoing links at the source node s that lie on a shortest route to the destination d. At each transit node t, the probability that a wavelength is available on one of $\min(l, n_{t,d})$ alternate outgoing links, given that a wavelength was available on an incoming link $L-1$ can be found to be

$$\alpha_1(\min(l, n_{t,d})) = 1 - \left(1 - \frac{\sum_{\overline{S}:\, |\overline{S}|<k} \pi(\overline{S}) \left(1 - \frac{S_1}{k(r-1)}\right)}{1 - \sum_{\overline{S}:\, |\overline{S}|\le k} \pi(\overline{S}) \frac{S_1}{k(r-1)}}\right)^{\min(l, n_{t,d})} \tag{8}$$

where r is the node degree.

In writing (7) and (8), we have used the approximation that the probability that an outgoing link at a node is available is independent of the availability of the alternate outgoing links at that node. However, as in Section II-B-1, we have *not* assumed that the probabilities of acquiring wavelengths on successive links of a session's path are independent. Our simulation results in Section V indicate that while our approximation is justified, the link-independence blocking approximation leads to rather poor results.

The probability of successfully establishing a circuit is then given by

$$P_{\text{succ}}(s, d) = \alpha_0(\min(l, n_{s,d})) \prod_{i=1}^{h(s,d)} \alpha_1\left(\min\left(l, n_{t(i),d}\right)\right)$$

where $t(i)$ is a transit node at hop i, and $h(s, d)$ is the number of hops on which a session is a transit session for a particular source–destination pair (s, d). For uniformly distributed destinations, the average probability of success P_{succ} for a new arrival can be found using (5), and the success probability $\overline{P}_{\text{succ}}$ of a random arrival (averaged over all trials, both new and reattempting) can be found using (6), with $P_{\text{succ}}(s, d)$ defined above.

III. Analysis for Torus Networks

In this section, we examine the $p \times p$ torus network, which consists of $N = p^2$ nodes arranged along the points of a 2-D grid with integer coordinates, with p nodes along each dimension. Two nodes (x_1, x_2) and (y_1, y_2) are connected by a bi-directional link if, and only if, for some $i = 1, 2$ we have $(x_i - y_i) \bmod p = 1$ and $x_j = y_j$ for $j \neq i$. In addition to these links, wraparound links connecting node $(x_1, 1)$ with node (x_1, p), and node $(1, x_2)$ with node (p, x_2) are also present.

A. Oblivious Routing

We first apply our analysis to oblivious routing, and study two different oblivious routing schemes: X–Y routing and Zig-Zag routing. Both schemes are shortest-hop routing schemes that differ in the method used to select the intermediate (transit) nodes along a route.

Given an outgoing link L, originating sessions that are emitted on L are defined as being of type $\tau = 0$. Transit sessions using L that arrive on an incoming link of a different dimension than L will be referred to as *bend* types ($\tau = 1$), while transit sessions using L that arrive over an incoming link of the same dimension as L will be referred to as *straight-through* types ($\tau = 2$).

1) X–Y Routing: X–Y routing is an oblivious routing scheme where a session follows a shortest route to its destination, first traversing all the links in one dimension (horizontal or vertical), and then traversing all the links in the other dimension (vertical or horizontal); the first dimension is selected as random at the source.

Applying Little's Theorem to the entire network and using the fact that the average number of wavelengths used is equal to the average number of active circuits in the system times the mean internodal distance, together with the symmetry of the torus network, the rate $\gamma(\tau)$ can be calculated to be

$$\gamma(\tau) = \begin{cases} \dfrac{\lambda}{4}, & \tau = 0 \text{ (originating)} \\ \dfrac{\lambda(p-1)}{4(p+1)}, & \tau = 1 \text{ (bend)} \\ \dfrac{\lambda \left\lceil \frac{p-2}{2} \right\rceil \left\lfloor \frac{p-2}{2} \right\rfloor}{4(p^2-1)}, & \tau = 2 \text{ (straight)} \end{cases} \tag{9}$$

and the rate $\gamma^*(\tau)$ for the auxiliary system can be found using (1). [The derivation of (9) can be found in the Appendix.]

Using X–Y routing, the path followed by a session with source s and destination d will make b bends along the way, where b is either 0 or 1, and will go straight through for a total of $i-b-1$ hops, where i is the shortest distance between s and d. The probability of successfully establishing a connection is given by

$$P_{\text{succ}}(s, d) = \alpha_0 \cdot \alpha_1^b \cdot \alpha_2^{i-b-1} \tag{10}$$

where α_0 is given by (3), and α_1 and α_2 are given by (4) with $M_1 = 2$ and $M_2 = 1$, respectively. For uniformly distributed destinations, the average probability of success for a new session can be calculated to be

$$P_{\text{succ}} = \begin{cases} \dfrac{\delta}{\alpha_1}\left[\left(1+2\alpha_1\left(\dfrac{1-\alpha_2^{(p-1)/2}}{1-\alpha_2}\right)\right)^2 - 1\right] & p \text{ odd} \\ \dfrac{\delta}{\alpha_1}\left[\left(1+2\alpha_1\left(\dfrac{1-\alpha_2^{p/2}}{1-\alpha_2}\right) + \alpha_1\left(\dfrac{1-\alpha_2^{(p/2)-1}}{1-\alpha_2}\right)\right)^2 - 1\right] & p \text{ even} \end{cases} \tag{11}$$

where $\delta = \alpha_0/(p^2-1)$. The average probability of success for a random connection attempt (either new or reattempting) can be written as

$$\overline{P}_{\text{succ}} = \begin{cases} \dfrac{\epsilon}{\alpha_1}\left[\left(1+\dfrac{2}{\alpha_1}\left(\dfrac{1-\alpha_2^{-((p-1)/2)}}{1-\alpha_2^{-1}}\right)\right)^2 - 1\right]^{-1} & p \text{ odd} \\ \dfrac{\epsilon}{\alpha_1}\left[\left(1+\dfrac{1}{\alpha_1}\left(\dfrac{1-\alpha_2^{-(p/2)}}{1-\alpha_2^{-1}}\right) + \dfrac{1}{\alpha_1}\left(\dfrac{1-\alpha_2^{1-(p/2)}}{1-\alpha_2^{11}}\right)\right)^2 - 1\right]^{-1} & p \text{ even} \end{cases} \tag{12}$$

where $\epsilon = \alpha_0(p^2-1)$.

2) Zig-Zag Routing: Zig-Zag routing is an oblivious routing scheme where one of the shortest routes from the source to the destination is selected with equal probability at the source. Note that the path may contain more than one turn (or bend), which is not the case for the X–Y routing discussed in the previous subsection.

To obtain the rate $\gamma(\tau)$ at which setup packets of type $\tau \in \{0, 1, 2\}$ leave an outgoing link L, we let Θ be the average probability that a setup packet is a straight-through transit type. If at all nodes both outgoing links lie on the shortest path, then Θ would be equal to $1/2$; however, in the torus network, the nodes that lie along the axis of the destination only have one preferred link. Therefore, when averaging over all nodes in the network, Θ is not exactly equal to $1/2$. (For example, in the 11×11 torus network, Θ works out to be 0.573.) The calculation of Θ is described in [29, App. A, Ch. 4]. The rates $\gamma(\tau)$ can be calculated using Little's Theorem and the symmetry of the torus to be

$$\gamma(\tau) = \begin{cases} \dfrac{\lambda}{4}, & \tau = 0 \text{ (originating)} \\ \dfrac{\lambda(\overline{h}-1)(1-\Theta)}{4}, & \tau = 1 \text{ (bend)} \\ \dfrac{\lambda(\overline{h}-1)\Theta}{4}, & \tau = 2 \text{ (straight)} \end{cases} \tag{13}$$

where $\overline{h}$ is the mean internodal distance of the torus and is given by

$$\overline{h} = \frac{p\left\lceil \dfrac{p^2-1}{2} \right\rceil}{p^2-1} \tag{14}$$

and the rates $\gamma^*(\tau)$ for the auxiliary system are calculated using (1). To derive (13), observe that we can use Little's Theorem to calculate the average number of wavelengths in use by a session of type τ in two ways: 1) as the product of the probability $((\gamma(\tau))/k)$ that a wavelength is in use by a session of type τ and the total number of wavelengths $4Nk$ in the system and 2) as the average number of sessions in the system times the average number of links on which a session is of type τ [e.g., for $\tau = 2$, this is $N\lambda\overline{X} \cdot (\overline{h}-1)\Theta$].

Assuming that a session at its source is at a distance of i hops from its destination, the average number of turns (or bend hops) that it will make is $(i-1)(1-\Theta)$, and the average number of links at which it will go straight through is $(i-1)\Theta$. The probability of successfully establishing a connection can, therefore, be approximated by

$$P_{\text{succ}}(i) \approx \alpha_0 \cdot \alpha_1^{(i-1)(1-\Theta)} \cdot \alpha_2^{(i-1)\Theta} \tag{15}$$

where α_0 is given by (3), and α_1 and α_2 are given by (4) with $M_1 = 2$ and $M_2 = 1$, respectively.

For uniformly distributed destinations, the average probability of success P_{succ} for a new session and the average probability of success $\overline{P}_{\text{succ}}$ for a random trial (either new or reattempting) can be written as

$$P_{\text{succ}} = \begin{cases} \dfrac{1}{p^2-1}\left[\displaystyle\sum_{i=1}^{(p-1)/2} 4iP_{\text{succ}}(i) + \sum_{i=(p+1)/2}^{p-1} 4(p-i)P_{\text{succ}}(i)\right] & p \text{ odd} \\ \dfrac{1}{p^2-1}\left[\displaystyle\sum_{i=1}^{(p/2)-1} 4iP_{\text{succ}}(i) + \sum_{i=(p/2)+1}^{p-1} 4(p-i)P_{\text{succ}}(i) + 2(p-1)P_{\text{succ}}\left(\frac{p}{2}\right) + P_{\text{succ}}(p)\right] & p \text{ even} \end{cases} \tag{16}$$

$$= \begin{cases} \dfrac{4\delta}{(1-\beta)^2}\left[1+\beta^p-\beta^{(p-1)/2}-\beta^{(p+1)/2}\right] & p \text{ odd} \\ \dfrac{4\delta}{(1-\beta)^2}\left[1+\beta^p-\beta^{(p/2)}-\dfrac{1}{2}\left(\beta^{(p/2)+1}+\beta^{(p/2)-1}\right)\right] & p \text{ even} \end{cases} \tag{17}$$

and

$$\overline{P}_{\text{succ}} = \begin{cases} (p^2-1) \left[\displaystyle\sum_{i=1}^{(p-1)/2} 4iP_{\text{succ}}^{-1}(i) + \sum_{i=(p+1)/2}^{p-1} 4(p-i)P_{\text{succ}}^{-1}(i) \right]^{-1} & p \text{ odd} \\ (p^2-1)\cdot \left[\displaystyle\sum_{i=1}^{(p/2)-1} 4iP_{\text{succ}}^{-1}(i) + \sum_{i=(p/2)+1}^{p-1} 4(p-i)P_{\text{succ}}^{-1}(i) + 2(p-1)P_{\text{succ}}^{-1}\left(\frac{p}{2}\right) + P_{\text{succ}}^{-1}(p) \right]^{-1} & p \text{ even} \end{cases} \tag{18}$$

$$= \begin{cases} \dfrac{\epsilon(\beta-1)^2}{4\beta^2} \left[1+\beta^{-p} - \beta^{(1-p)/2} - \beta^{(-1-p)/2}\right]^{-1} & p \text{ odd} \\ \dfrac{\epsilon(\beta-1)^2}{4\beta^2} \left[1+\beta^{-p} - \beta^{-p/2} - \frac{1}{2}\left(\beta^{-1-(p/2)} + \beta^{1-(p/2)}\right)\right]^{-1} & p \text{ even} \end{cases} \tag{19}$$

where $\beta = \alpha_1^{(1-\Theta)}\alpha_2^{\Theta}$, $\delta = \alpha_0/(p^2-1)$, and $\epsilon = \alpha_0(p^2-1)$.

B. Adaptive Routing

We now find the success probability for the torus network when adaptive routing is used. Following the analysis of Section II-B-2, we define only two types of sessions: originating sessions (type $\tau = 0$) and transit sessions (type $\tau = 1$). Note that for the torus network, there are at most two outgoing links at a node that lie a shortest route to the destination.

Using Little's Theorem and the symmetry of the torus, the rates $\gamma(\tau)$ can be calculated as

$$\gamma(\tau) = \begin{cases} \dfrac{\lambda}{4}, & \tau = 0 \text{ (originating)} \\ \dfrac{\lambda(\overline{h}-1)}{4}, & \tau = 1 \text{ (transit)} \end{cases} \tag{20}$$

where $\overline{h}$ is the mean internodal distance of the torus and is given by (14).

In the torus network, at a distance i, $i \le \lfloor p/2 \rfloor$ from the destination, there are four nodes that have only one outgoing link along a shortest route, and there are $4i - 4$ nodes that have two outgoing links that lie on the shortest route. All other nodes that are at a distance i, $i > \lfloor p/2 \rfloor$ have two outgoing links that lie on the shortest route. Assuming that a session at its source is at a distance i hops from its destination, and the session is allowed to try at most two outgoing links per hop, the probability of successfully establishing a connection is given by

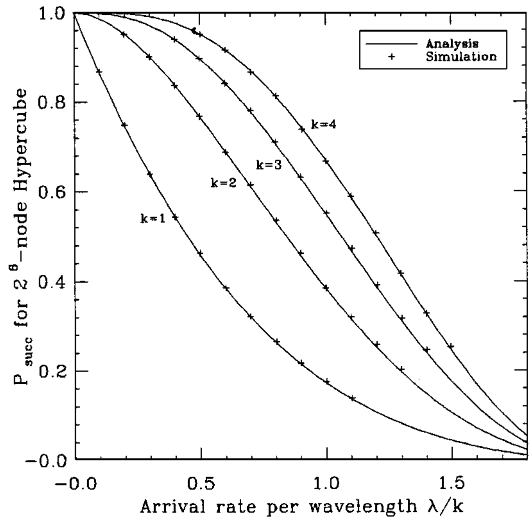


Fig. 2. Analytical and simulation results for P_{succ} versus the arrival rate per wavelength λ/k for a 2^6-node hypercube network using oblivious routing.

$$P_{\text{succ}}(i) = \begin{cases} (\alpha_0(1) + (i-1)\alpha_0(2)) \left(\dfrac{1}{i!}\right) \cdot \displaystyle\prod_{b=1}^{i-1} (\alpha_1(1) + (b-1)\alpha_1(2)), & i \le \left\lceil \dfrac{p-1}{2} \right\rceil \\ \alpha_0(2)\alpha_1(2)^{i-\lceil (p+1)/2 \rceil} \left(\frac{1}{(\lceil (p-1)/2 \rceil)!}\right) \cdot \displaystyle\prod_{b=1}^{\lceil (p-1)/2 \rceil} (\alpha_1(1) + (b-1)\alpha_1(2)), & i > \left\lceil \dfrac{p-1}{2} \right\rceil \end{cases} \tag{21}$$

where $\alpha_0(i)$, $i = 1, 2$ is given by (7), and $\alpha_1(i)$, $i = 1, 2$ is given by (8) with $M_1 = 3$. In writing (21) for node pairs at a distance i, $i \le \lfloor p/2 \rfloor$, we used the average probability that originating and transit links are available.

For uniformly distributed destinations, the average probability of success for a new arrival P_{succ} and a random arrival (averaged over all trials, both new and reattempting) $\overline{P}_{\text{succ}}$ can be found using (16) and (18), respectively, where $P_{\text{succ}}(i)$ is defined in (21).

IV. Analysis for Hypercube Networks

In this section, we turn our attention to the 2^r-node hypercube network, where each node can be represented by a binary string $(x_1, x_2, \ldots, x_r)$, and two nodes are connected via a bidirectional link if their binary representations differ in only one bit. Given an outgoing link L, we observe that due to symmetry, all the other incoming links are of the same spatial type. Thus, originating sessions that are emitted on L are defined as being of type $\tau = 0$, while all transit sessions using L are defined as being of type $\tau = 1$.

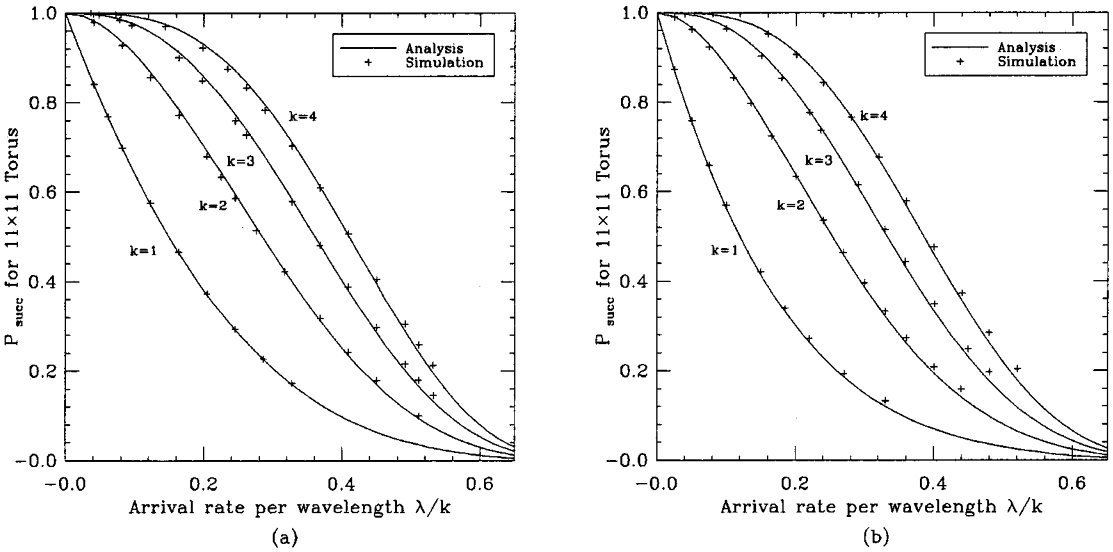


Fig. 3. Analytical and simulation results for P_{succ} versus the arrival rate per wavelength λ/k for an 11×11 torus using oblivious routing: (a) P_{succ} with X–Y routing and (b) P_{succ} with Zig-Zag routing.

A. Oblivious Routing

We first find the success probability for oblivious routing in the hypercube network, where a shortest route is chosen at random at the source.

The rates $\gamma(\tau)$ at which setup packets of type τ are emitted on a link can be found to be

$$\gamma(\tau) = \begin{cases} \dfrac{\lambda}{r}, & \tau = 0 \text{ (originating)} \\ \dfrac{\lambda\left[(r-2)2^{r-1}+1\right]}{r(2^r-1)}, & \tau = 1 \text{ (transit)} \end{cases} \tag{22}$$

and $\gamma^*(\tau)$ can be found using (1).

Assuming the source is at a distance of i hops from its destination, the probability of successfully establishing a connection is given by

$$P_{\text{succ}}(i) = \alpha_0 \cdot \alpha_1^{i-1} \tag{23}$$

where α_0 is given by (3) and α_1 is given by (4) with $M_1 = r-1$.

For uniformly distributed destinations, the average probability of success for a new arrival can be written as

$$P_{\text{succ}} = \frac{\alpha_0}{\alpha_1(2^r-1)}\left[(1+\alpha_1)^r - 1\right] \tag{24}$$

where we have used the fact that there are $\binom{r}{i}$ nodes at a distance i from a given node. The average probability of success for a random connection attempt (either new or reattempting) can be written as

$$\overline{P}_{\text{succ}} = \frac{\alpha_0(2^r-1)}{\alpha_1}\left[\left(1+\frac{1}{\alpha_1}\right)^r - 1\right]^{-1}. \tag{25}$$

B. Adaptive Routing

To find the success probability for a hypercube network using adaptive routing, we again consider the adaptive routing scheme of Section II-B-2 and note that, in the hypercube network, a node that is i hops away from the destination has i outgoing links lying along a shortest route to the destination. We let l, $l \leq r$, be the maximum number of outgoing links that may be tried at any hop.

The rates $\gamma(\tau)$ at which setup packets of type τ are emitted on a link are given by (22). Assuming the source is at a distance i hops from its destination, the probability of successfully establishing a connection is given by

$$P_{\text{succ}}(i) = \alpha_0(\min(l, i)) \prod_{j=1}^{i-1} \alpha_1(\min(l, j)) \tag{26}$$

where $\alpha_0(\min(l, i))$, $i = 1, \ldots, r$, is given by (7), and $\alpha_1(\min(l, j))$, $j = 1, \ldots, r$ is given by (8). The probabilities P_{succ} and $\overline{P}_{\text{succ}}$ can then be found using (24) and (25), respectively.

V. Analytical and Simulation Results

In this section, we present our analytical and simulation results for oblivious and adaptive routing in the torus and hypercube networks. We first demonstrate the accuracy of our analysis by comparing the probabilities of success obtained from our analysis with those obtained from simulations for oblivious and adaptive routing in both the torus and hypercube networks. Next, we show that the link independence blocking assumption used in other analyses fares poorly for both oblivious and adaptive routing. We then compare the benefits obtained by increasing the number of wavelengths (larger k) with those obtained by increasing the routing flexibility (larger l), and demonstrate that an interesting tradeoff exits between the two, with important implications for network provisioning in all-optical networks.

In Figs. 2 and 3, we compare the success probability P_{succ} predicted by our analysis with those obtained from simulations for oblivious routing in the hypercube and torus networks, respectively, while in Fig. 4 we present results for adaptive routing; comparisons for $\overline{P}_{\text{succ}}$ can be found in [29, Sec. 4.5, Ch. 4]. We observe that in all the figures, there is close agreement between the simulations and the analytically predicted

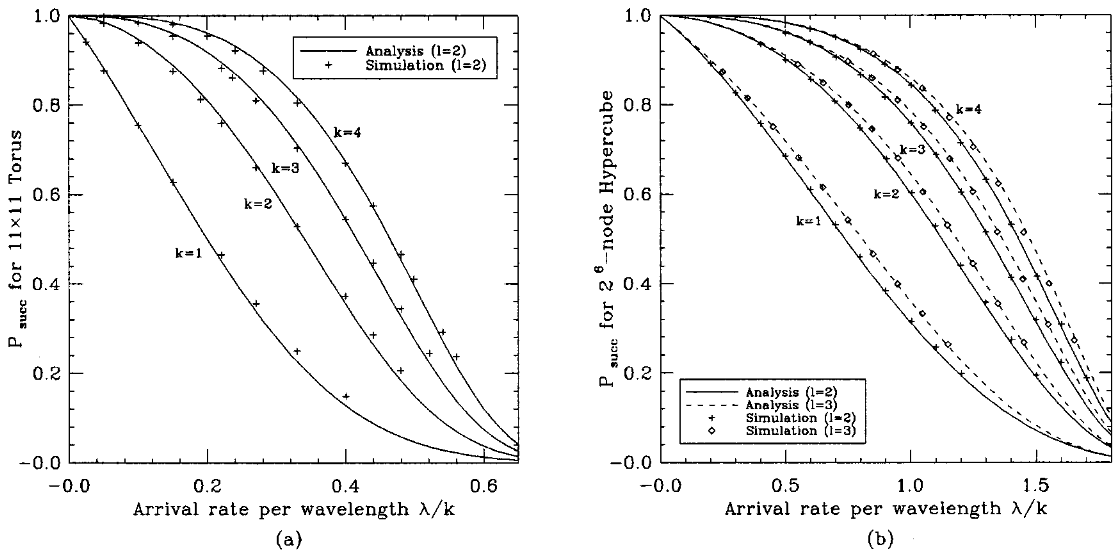


Fig. 4. Analytical and simulation results for $P_{\rm succ}$ versus the arrival rate per wavelength λ/k for an 11×11 torus and a 2^6-node hypercube network using adaptive routing: (a) $P_{\rm succ}$ for the torus network and (b) $P_{\rm succ}$ for the hypercube network.

values over the entire range of applicable input rates, which is a significant improvement over previous analyses. Despite its accuracy, our analysis is considerably simpler than the analyses available in the literature, and its computational requirements are modest, allowing it to scale easily for large k.

Our analysis shows that for a setup packet in the torus network, the probability α_2 of successfully obtaining an output wavelength when going straight through is larger than the probability α_1 of successfully obtaining an output wavelength when making a bend or turn. (We illustrate this for an 11×11 torus network in Table I for both X–Y and Zig-Zag routing and for the number of wavelengths $k = 1, 2$, and 4.) Recall from Section III-A that in X–Y routing a source–destination path may contain at most one bend hop [hence b, the exponent of α_1 in (10), is either 0 or 1] and $i - b - 1$ straight-through hops (where i is the shortest distance between the source and destination). Since bend hops reduce α_2 and straight-through hops affect reduce α_1, the probability α_2 will be larger than α_1. For Zig-Zag routing, even though paths may contain multiple turns, there are still, on average, more straight-through hops than there are bend hops. This is because once a setup packet reaches a node along the x or y axis of the destination, it must continue along that axis toward the destination (so all future hops must be straight-through hops). As with X–Y routing, this translates to $\alpha_2 > \alpha_1$. However, the difference between the two is larger for X–Y routing than it is for Zig-Zag routing for the reasons explained before (see also Table I).

Since the probability of success when making a turn is smaller than the probability of success when going straight through, the probability of success for X–Y routing (which has minimal turns) is expected to be larger than the probability of success for Zig-Zag routing (which has on average more turns than X–Y routing). This result is confirmed by comparing the curves in Fig. 3(a) with those in Fig. 3(b). These nonobvious results correspond to significant performance differentials. For instance, with $P_{\rm succ} = 0.8$ and $k = 1$, X–Y routing gives a maximum throughput that is 30% greater than that for Zig-Zag routing.

TABLE I

ANALYTICAL RESULTS FOR PROBABILITIES α_1 AND α_2, $k = 1, 2$, AND 4, OF SUCCESSFULLY OBTAINING AN OUTPUT WAVELENGTH WHEN GOING STRAIGHT THROUGH AND MAKING A BEND, RESPECTIVELY, FOR AN 11×11 TORUS AND VARIOUS VALUES OF THE ARRIVAL RATE PER WAVELENGTH λ/k

Torus (X-Y Routing)						
	k=1		k=2		k=4	
λ/k	α_1	α_2	α_1	α_2	α_1	α_2
0.0375	0.9522	0.9822	0.9955	0.9983	0.9999	1
0.0750	0.9039	0.9631	0.9833	0.9936	0.9993	0.9997
0.1500	0.8063	0.9203	0.9408	0.9756	0.9919	0.9966
0.2250	0.7072	0.8701	0.8784	0.9461	0.9704	0.9869
0.3000	0.6065	0.8103	0.7989	0.9031	0.9302	0.9664
0.3750	0.5041	0.7381	0.7031	0.8432	0.8674	0.9300
0.4500	0.4000	0.6489	0.5911	0.7607	0.7775	0.8698
0.5250	0.2942	0.5361	0.4618	0.6463	0.6537	0.7724
0.6000	0.1867	0.3889	0.3128	0.4837	0.4842	0.6124
0.6750	0.0773	0.1885	0.1397	0.2434	0.2434	0.3346
Torus (Zig-Zag Routing)						
	k=1		k=2		k=4	
λ/k	α_1	α_2	α_1	α_2	α_1	α_2
0.0375	0.9571	0.9719	0.9960	0.9974	0.9999	1
0.0750	0.9133	0.9424	0.9850	0.9900	0.9993	0.9996
0.1500	0.8234	0.8787	0.9460	0.9629	0.9925	0.9949
0.2250	0.7301	0.8078	0.8879	0.9202	0.9727	0.9806
0.3000	0.6331	0.7283	0.8125	0.8611	0.9350	0.9519
0.3750	0.5323	0.6388	0.7200	0.7837	0.8750	0.9034
0.4500	0.4275	0.5370	0.6098	0.6845	0.7877	0.8283
0.5250	0.3183	0.4204	0.4801	0.5580	0.6655	0.7156
0.6000	0.2045	0.2854	0.3279	0.3962	0.4955	0.5468
0.6750	0.0858	0.1272	0.1476	0.1863	0.2504	0.2844

Similar results hold for the entire range of network loads and all values of k that we examined, with a decrease in improvement as the load increases.

In Fig. 5, we show the results that would have been obtained if the independence blocking assumption was used, together with our analysis and simulations, for both oblivious and adaptive routing. Note that the link independence blocking approximation gives very poor results for the torus network, and there is only a slight improvement as k increases. This is because the torus is a sparse topology with small node degree, and sessions are not mixed well, resulting in a high correlation between the wavelengths used on successive links on a session's path. Observe also that while the link independence blocking assumption

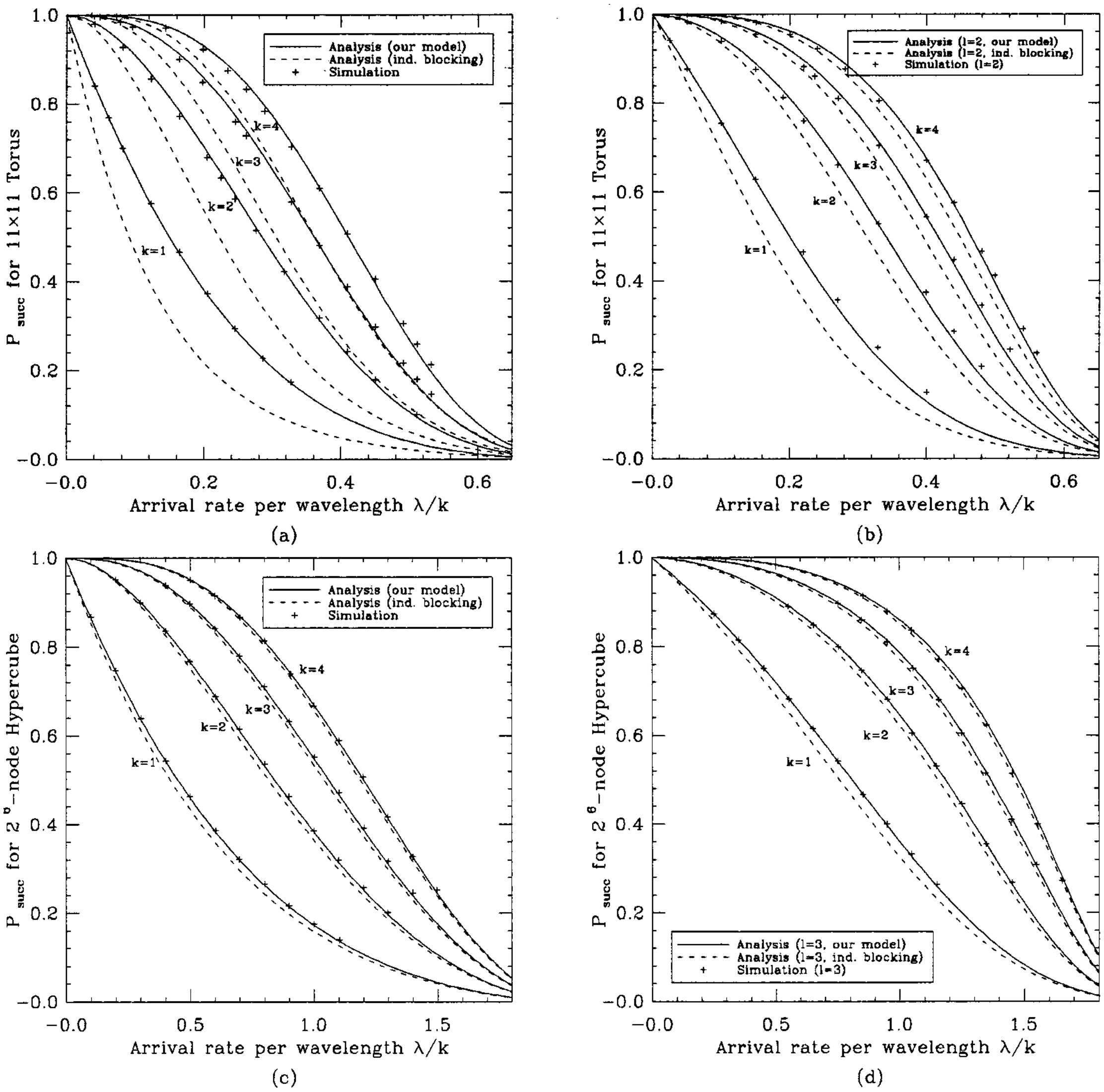


Fig. 5. Analytical and simulation results, together with results that use the link independence blocking assumption for $P_{\rm succ}$ versus the arrival rate per wavelength λ/k for an 11×11 torus network and a 2^6-node hypercube network: (a) oblivious (X–Y) routing; (b) adaptive routing with $l = 2$ for the torus; (c) oblivious (random) routing; and (d) adaptive routing with $l = 3$ for the hypercube.

is a poor approximation, the approximation that at a particular hop, the occupancy probabilities of the alternate links are independent, is a very good one.

To evaluate the throughput gains obtained by varying k and l, we define $\lambda(P_{\rm succ}, k, l)$ as the throughput per node per wavelength in a system with k wavelengths and routing flexibility l, when the probability of success is equal to $P_{\rm succ}$. Similarly, we define $P_{\rm succ}(\lambda, k, l)$ to be the probability of success in a system with k wavelengths and routing flexibility l, when the arrival rate per node per wavelength is equal to λ/k. To compare the performance of systems with varying k and l, we define the *incremental per-wavelength throughput gain* $\Delta\lambda(k_1, l_1; k_2, l_2)$ of a system with k_2 wavelengths and a choice of l_2 links per hop, over a system with k_1 wavelengths and a choice of l_1 links per hop, for a given $P_{\rm succ}$, to be

$$\Delta\lambda(k_1, l_1; k_2, l_2) = \frac{\lambda(P_{\rm succ}, k_2, l_2) - \lambda(P_{\rm succ}, k_1, l_1)}{\lambda(P_{\rm succ}, k_1, l_1)} \times 100\%. \quad (27)$$

We also define the *incremental probability of success gain* $\Delta P_{\rm succ}(k_1, l_1; k_2, l_2)$ of a system with k_2 wavelengths and a choice of l_2 links per hop over a system with k_1 wavelengths and a choice of l_1 links per hop, for a given λ/k, to be

$$\Delta P_{\rm succ}(k_1, l_1; k_2, l_2) = \frac{P_{\rm succ}(\lambda, k_2, l_2) - P_{\rm succ}(\lambda, k_1, l_1)}{P_{\rm succ}(\lambda, k_1, l_1)} \times 100\%. \quad (28)$$

The throughput and probability of success gains measure the degree of improvement that a full-wavelength translation system with k_2 wavelengths and a choice of l_2 outgoing links per hop provides over a similar system with k_1 wavelengths and a choice of l_1 links per hop.

In Fig. 6, we illustrate the analytically predicted probability of success $P_{\rm succ}$ versus the arrival rate per wavelength λ/k, for k ranging from 1 to 16, for both the torus and hypercube networks. In Tables II and III, we show the per-wavelength incremental throughput gains for two values of $P_{\rm succ}$ for both oblivious and adaptive routing, for an 11×11 torus network and for a

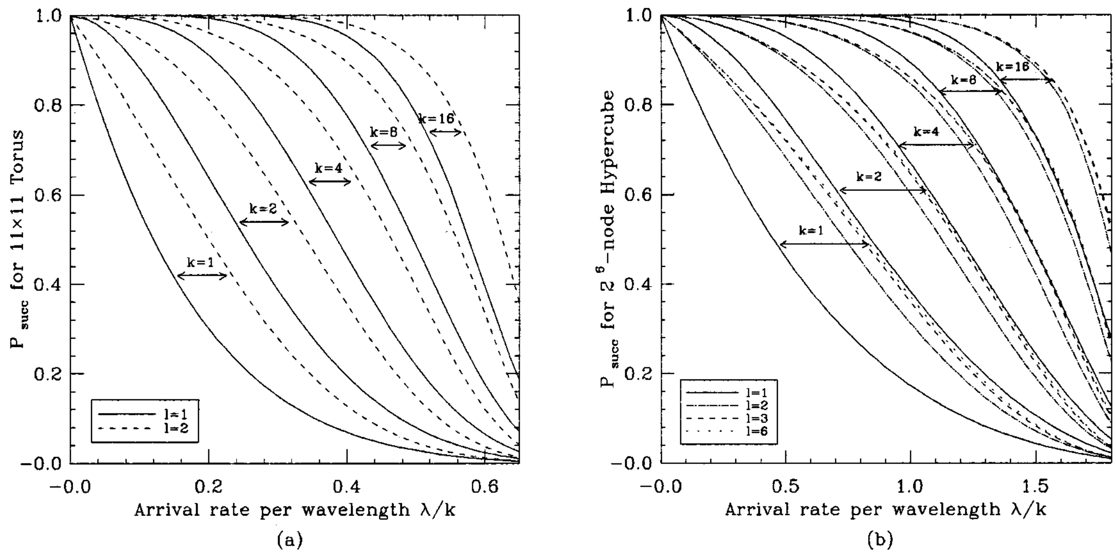


Fig. 6. The probability of success P_{succ} for (a) an 11×11 torus and (b) 2^6-node hypercube network, for k varying from 1 to 16.

TABLE II
INCREMENTAL PER-WAVELENGTH THROUGHPUT GAINS $\Delta\lambda(k_1, l_1; k_2, l_2)$ FOR AN 11×11 TORUS AND A 2^6-NODE HYPERCUBE NETWORK FOR OBLIVIOUS ROUTING WITH P_{succ} EQUAL TO 0.8 AND 0.5. THE ROW ELEMENTS FROM LEFT TO RIGHT CORRESPOND TO INCREASING THE NUMBER OF WAVELENGTHS k, WHILE HOLDING THE ROUTING FLEXIBILITY l CONSTANT

Torus (X-Y Routing)				
P_{succ}	$\Delta\lambda(1,1;2,1)$	$\Delta\lambda(2,1;4,1)$	$\Delta\lambda(4,1;8,1)$	$\Delta\lambda(8,1;16,1)$
0.8	200%	92%	43%	23%
0.5	87%	46%	26%	15%
Torus (Zig-Zag Routing)				
P_{succ}	$\Delta\lambda(1,1;2,1)$	$\Delta\lambda(2,1;4,1)$	$\Delta\lambda(4,1;8,1)$	$\Delta\lambda(8,1;16,1)$
0.8	267%	91%	48%	26%
0.5	108%	54%	28%	16%
Hypercube (Random Routing)				
P_{succ}	$\Delta\lambda(1,1;2,1)$	$\Delta\lambda(2,1;4,1)$	$\Delta\lambda(4,1;8,1)$	$\Delta\lambda(8,1;16,1)$
0.8	195%	81%	41%	23%
0.5	87%	42%	23%	13%

TABLE III
INCREMENTAL PER-WAVELENGTH THROUGHPUT GAINS $\Delta\lambda(k_1, l_1; k_2, l_2)$ FOR AN 11×11 TORUS AND A 2^6-NODE HYPERCUBE NETWORK FOR ADAPTIVE ROUTING WITH P_{succ} EQUAL TO 0.8 AND 0.5. THE TOP-TO-BOTTOM ELEMENTS OF A COLUMN (WITHIN EACH NETWORK) CORRESPOND TO INCREASING THE ROUTING FLEXIBILITY l, WHILE HOLDING THE NUMBER OF WAVELENGTHS k CONSTANT

Torus				
P_{succ}	$\Delta\lambda(1,1;1,2)$	$\Delta\lambda(2,1;2,2)$	$\Delta\lambda(4,1;4,2)$	$\Delta\lambda(8,1;8,2)$
0.8	61%	32%	19%	18%
0.5	67%	35%	21%	13%
Hypercube				
P_{succ}	$\Delta\lambda(1,1;1,2)$	$\Delta\lambda(2,1;2,2)$	$\Delta\lambda(4,1;4,2)$	$\Delta\lambda(8,1;8,2)$
0.8	121%	57%	32%	19%
0.5	64%	33%	19%	11%
P_{succ}	$\Delta\lambda(1,2;1,3)$	$\Delta\lambda(2,2;2,3)$	$\Delta\lambda(4,2;4,3)$	$\Delta\lambda(8,2;8,3)$
0.8	10%	5.6%	3.4%	2.1%
0.5	9.3%	5.3%	3.1%	1.8%
P_{succ}	$\Delta\lambda(1,3;1,4)$	$\Delta\lambda(2,3;2,4)$	$\Delta\lambda(4,3;4,4)$	$\Delta\lambda(8,3;8,4)$
0.8	0.8%	0.5%	0.3%	0.2%
0.5	1.6%	0.8%	0.5%	0.3%

2^6-node hypercube network. In Table II, the row elements from left to right correspond to increasing k, while holding l constant. For example, in the torus network with X–Y routing and $P_{\text{succ}} = 0.8$, using a full-wavelength translation system with two wavelengths per link achieves a 92%; gain in throughput per wavelength over a system with one wavelength per link (i.e., with no-wavelength translation). In Table III, the top-to-bottom elements of a column (within each network) correspond to increasing l, while holding k constant. For example, in the hypercube network with $P_{\text{succ}} = 0.8$, a system with two wavelengths per link and adaptive routing (with $l = 2$) achieves a 57% gain in the throughput per wavelength over a system with two wavelengths per link and oblivious routing ($l = 1$).

As is evident from Fig. 6 and Table II, for a given P_{succ} and fixed l, the throughput per wavelength increases with increasing k. In other words, the throughput per link (and the network throughput) increases superlinearly with k. The linear part of the increase in throughput is because of the increase in capacity, while the superlinear part of the increase is due to more efficient use of that capacity because of the greater flexibility in establishing a circuit when a larger number of wavelengths is available. The incremental gain in achievable throughput per wavelength for a given l, $\Delta\lambda(k_1, l; k_2, l)$, however, decreases rapidly with increasing k. This result holds for both oblivious and adaptive routing, and is in agreement with the results for oblivious routing presented in [30] and [8]. Similarly, the incremental throughput gain for a given k, $\Delta\lambda(k, l_1; k, l_2)$, decreases rapidly with increasing l. If we fix l_1 and l_2, and increase k, the incremental gain decreases, suggesting that the performance improvement for adaptive routing is tightly coupled with the number of wavelengths, and that the benefits of alternate routing are not as significant when the number of wavelengths k is large. [For example, in the 2^6-hypercube network, $\Delta\lambda(2, 1; 2, 2) = 57\%$, while $\Delta\lambda(4, 1; 4, 2) = 32\%$.] In all our throughput calculations, we have neglected the overhead due to circuits being partially established and blocked.

Another interesting feature of adaptive routing with wavelength translation is that the per-wavelength throughput for fixed k and increasing l, appears to saturate at or near the per-wavelength throughput of a system using oblivious routing with wavelength translation over twice as many wavelengths. In Fig. 7, we plot the incremental throughput gain for the hyper-

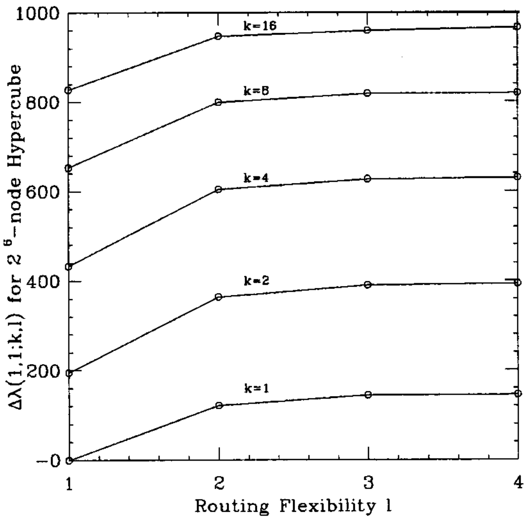


Fig. 7. The incremental throughput gain $\Delta\lambda(1,\, 1;\, k,\, l)$ for $k = 1,\, 2,\, 4,\, 8,\, 16$ and $l = 1,\, 2,\, 3,\, 4$ for a 2^6-node hypercube network with $P_{\rm succ} = 0.8$.

cube network when $P_{\rm succ} = 0.8$ and the number of wavelengths k ranges from 1 to 16 and the routing flexibility l ranges from 1 to 4 [i.e., we plot $\Delta\lambda(1,\, 1;\, k,\, l)$ for $k = 1,\, 2,\, 4,\, 8,\, 16$ and $l = 1,\, 2,\, 3,\, 4$]. As can be seen in Fig. 7, the largest increase in incremental throughput gain occurs when the routing flexibility increases from $l = 1$ to $l = 2$, regardless of the number of wavelengths. Furthermore, this gain obtained by increasing the routing flexibility from $l = 1$ to $l = 2$, with fixed k, approaches the gain obtained by doubling the number of wavelengths to $2k$, with $l = 1$. For example, the incremental throughput gain for $k = 8$ and $l = 2$ is within 3% of the incremental throughput gain for $k = 16$ and $l = 1$.

The above discussion leads to some interesting design options when building an all-optical network. For instance, since the per-wavelength throughput gain saturates quickly with increasing k, simply building a network in which every node can translate between k wavelengths may not be the most efficient option. Instead, it may be preferable to build a network in which every node consists of k/n simpler switching elements operating in parallel (each switching between a nonintersecting subset of n wavelengths) that achieves performance comparable to that of the k-wavelength system at a much lower cost. For example, using the switch implementation of Yoo and Bala [31] (to be specific), the component cost (in terms of elementary 2×2 switches) is $kr\log(nr) - ((nr)/2)$, whereas the increase in success probability is negligible as n increases beyond some point. This suggests that a network designer may initially choose to build the network with nodes that have a small number of parallel channels, with n wavelengths per channel. As network traffic grows, the designer may expand the nodes by adding more parallel channels. Better yet, instead of increasing the number of channels per link at every network node, the designer may focus on the routing algorithms and may choose to increase the routing flexibility to obtain equivalent performance at no extra hardware cost. For instance, the designer may simply increase the number of outgoing links that may be tried at each hop. Observe, however, that the routing flexibility is limited by the network topology and also is a function of the switch architecture. Our results emphasize the need for network designers to investigate the tradeoffs between wavelength translation, routing flexibility, and hardware cost when designing future optical networks.

VI. Conclusion

We presented a new general analysis for oblivious and adaptive routing in all-optical regular networks with wavelength translation that is intuitive, simple, computationally inexpensive, and the first to consider adaptive routing with wavelength translation. Our analysis does not use the link independence blocking assumption, and is more accurate than previous analyses (for oblivious routing) over a wider range of network loads. We verified our analysis for the hypercube and torus topologies, and found that although the throughput per wavelength increases with an increase in the number of wavelengths k, and an increase in the number of link choices l for adaptive routing, this increase saturates quickly. We showed that for the topologies considered, the performance of a system using adaptive routing with only one alternate link per hop, approaches that of a system using oblivious routing with twice as many wavelengths per link. We also showed that for the torus network, X–Y routing performs better than Zig-Zag routing. These observations lead to some interesting possibilities for provisioning an all-optical network from a performance-cost perspective.

Appendix
Derivation of (9)

In this appendix, we derive (9) of Section III-A-1. For X–Y routing, the *routing tag* of a session with source node $s = (s_1,\, s_2)$ and destination node $d = (d_1,\, d_2)$, is defined as $(t_1,\, t_2)$, where

$$t_j = \begin{cases} d_j - s_j, & \text{if } |d_j - s_j| \le \left\lfloor \dfrac{p}{2} \right\rfloor \\ d_j - s_j - p \cdot \text{sgn}(d_j - s_j), & \text{if } |d_j - s_j| > \left\lfloor \dfrac{p}{2} \right\rfloor \end{cases}$$

for all $j \in 1,\, 2$, and where $\text{sgn}(x)$ is the signum function, which is equal to $+1$ if $x \ge 0$, and equal to -1, otherwise.

Recall that in our routing scheme, destinations are distributed uniformly over all nodes (excluding the source node) and blocked sessions are not dropped, but reinserted back into the input stream. Thus, wavelength utilization is uniform across all wavelengths of the network, and the probabilities $q(\tau)$ that a wavelength is in use by a session of type τ, $\tau = 0,\, 1,\, 2$ or is idle $\tau = 3$ can be obtained simply by counting all possible ways in which a wavelength can be in use by such a session, and normalizing appropriately.

Thus, the probability $q(\tau)$ that a wavelength on an outgoing link L is used by a bend session ($\tau = 1$) is given by

$$q(1) = C \sum_{\substack{t_2 = -\lfloor p/2 \rfloor \\ t_2 \ne 0}}^{\lfloor p/2 \rfloor} \sum_{\substack{t_1 = -\lfloor p/2 \rfloor \\ t_1 \ne 0}}^{\lfloor p/2 \rfloor} 1 = C(p-1)^2. \qquad \text{(A.1)}$$

Similarly, the probability $q(\tau)$ that a wavelength on an outgoing line L is used by a straight-through session ($\tau = 2$) is given by

$$q(2) = C\left[\sum_{\substack{t_2=-\lfloor p/2\rfloor\\ t_2\neq 0}}^{\lfloor p/2\rfloor}\sum_{\substack{t_1=-\lfloor p/2\rfloor\\ t_1\neq 0}}^{\lfloor p/2\rfloor}(|t_2|+|t_1|-2) + 2\sum_{\substack{t_2=-\lfloor p/2\rfloor\\ t_2\neq 0}}^{\lfloor p/2\rfloor}(|t_2|-1)\right] = C\cdot\left\lceil\frac{p-2}{2}\right\rceil\left\lfloor\frac{p-2}{2}\right\rfloor. \quad \text{(A.2)}$$

The remaining probability $q(\tau)$ that a wavelength on link L is used by a session originating at a node ($\tau = 0$) is given by

$$q(0) = C(p^2-1). \quad \text{(A.3)}$$

Finally, by simple application of Little's Theorem, where we equate the average number of wavelengths used $4Nk(1-q(3))$ to the product of the average number of active sessions in the system $N\lambda\overline{X}$ times the mean internodal distance $\overline{h}$, we obtain

$$1-q(3) = \frac{\lambda\overline{X}\overline{h}}{4k}. \quad \text{(A.4)}$$

Equations (A.1)–(A.4), together with the condition

$$\sum_{\tau=0}^{3} q_i = 1$$

can be used to calculate C, so that we finally get

$$q(\tau) = \begin{cases} \dfrac{\lambda\overline{X}\overline{h}}{4k}\cdot\dfrac{p^2-1}{p\cdot\left\lceil\dfrac{p^2-1}{2}\right\rceil}, & \tau = 0 \\ \dfrac{\lambda\overline{X}\overline{h}}{4k}\cdot\dfrac{(p-1)^2}{p\cdot\left\lceil\dfrac{p^2-1}{2}\right\rceil}, & \tau = 1 \\ \dfrac{\lambda\overline{X}\overline{h}}{4k}\cdot\dfrac{\left\lceil\dfrac{p-2}{2}\right\rceil\left\lfloor\dfrac{p-2}{2}\right\rfloor}{p\cdot\left\lceil\dfrac{p^2-1}{2}\right\rceil}, & \tau = 2 \end{cases} \quad \text{(A.5)}$$

where $\lfloor x\rfloor$ is the largest integer less than or equal to x, and $\lceil x\rceil$ is the smallest integer greater than or equal to x. Equations (A.2)–(A.5) hold for both p odd and p even.

Now, applying Little's Theorem to the system that consists only of sessions of type τ, and equating the average number of wavelengths used by type τ sessions to the product of the total arrival rate of type τ sessions into the system and the session holding time, we obtain

$$4Nkq(\tau) = 4Nk\frac{\gamma(\tau)}{k}\cdot\overline{X}$$

which gives

$$\gamma(\tau) = \frac{k\cdot q(\tau)}{\overline{X}}$$

from which (9) readily follows.

REFERENCES

[1] A. Acampora, "The scalable lightwave network," *IEEE Commun. Mag.*, vol. 32, pp. 36–42, Dec. 1994.

[2] R. Barry and P. Humblet, "Models of blocking probability in all-optical networks with and without wavelength changers," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 858–867, June 1996.

[3] M. Kovačević and A. Acampora, "Benefits of wavelength translation in all-optical clear-channel networks," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 868–880, June 1996.

[4] J. P. Lang, V. Sharma, and E. A. Varvarigos, "A new analysis of wavelength translation in regular WDM networks," in *Proc. MPPOI'98*, June 1998, pp. 131–139.

[5] J. Yates, J. Lacey, D. Everitt, and M. Summerfield, "Limited-range wavelength translation in all-optical networks," in *Proc. IEEE INFOCOM'96*, vol. 3, Mar. 1996, pp. 954–961.

[6] N. Wauters and P. Demeester, "Wavelength translation in optical multi-wavelength multi-fiber transport networks," *Int. J. Optoelectron.*, vol. 11, pp. 53–70, Jan.–Feb. 1997.

[7] R. Ramaswami and G. H. Sasaki, "Multiwavelength optical networks with limited wavelength conversion," in *Proc. IEEE INFOCOM'97*, vol. 2, Apr. 1997, pp. 89–98.

[8] V. Sharma and E. A. Varvarigos, "An analysis of limited wavelength translation in regular all-optical WDM networks," in *Proc. IEEE INFOCOM'98*, vol. 2, Mar.–Apr. 1998, pp. 893–901.

[9] H. Harai, M. Murata, and H. Miyahara, "Performance analysis of wavelength assignment policies in all-optical networks with limited-range wavelength conversion," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1051–1060, Sep. 1998.

[10] R. Barry, "Wavelength routing for all-optical networks," Ph.D. dissertation, MIT, Cambridge, MA, 1993.

[11] N. Nagatsu, Y. Hamazumi, and K. Sato, "Number of wavelengths required for constructing large-scale optical path networks," *Electron. Commun. Jpn.*, pt. 1, vol. 78, pp. 1–10, Sept. 1995.

[12] A. Birman, "Computing approximate blocking probabilities for a class of all-optical networks," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 852–857, Jun. 1996.

[13] S. Baroni, P. Bayvel, and J. Midwinter, "Wavelength requirements in dense wavelength-routed optical transport networks with variable physical connectivity," *Electron. Lett.*, vol. 32, pp. 575–576, Mar. 1996.

[14] S. Subramaniam and R. Barry, "Wavelength assignment in fixed routing WDM networks," in *Proc. ICC'97*, vol. 1, Jun. 1997, pp. 406–410.

[15] K. N. Sivarajan and R. Ramaswami, "Lightwave networks based on de Bruijn graphs," *IEEE/ACM Trans. Networking*, vol. 2, pp. 70–79, Feb. 1994.

[16] Y. Zhu, G. Rouskas, and H. Perros, "Blocking in wavelength routing networks, Part 1: The single path case," in *Proc. IEEE INFOCOM'99*, vol. 1, Mar. 1999, pp. 321–328.

[17] A. Karasan and E. Ayanoglu, "Effects of wavelength routing and wavelength selection algorithms on wavelength conversion gain in WDM optical networks," *IEEE/ACM Trans. Networking*, vol. 6, pp. 186–196, Apr. 1998.

[18] A. Mokhtar and M. Azizoglu, "A new analysis of wavelength translation in regular WDM networks," *IEEE/ACM Trans. Networking*, vol. 6, pp. 197–206, Apr. 1998.

[19] H. Harai, M. Murata, and H. Miyahara, "Performance of alternate routing methods in all-optical switching networks," in *Proc. IEEE INFOCOM'97*, vol. 2, Apr. 1997, pp. 516–524.

[20] F. Kelly, "Blocking probabilities in large circuit switched networks," in *Adv. Appl. Prob.*, 1986, vol. 18, pp. 473–505.

[21] S.-P. Chung and K. Ross, "Reduced load approximations for multirate loss network," *IEEE Trans. Commun.*, vol. 41, pp. 1222–1231, Aug. 1993.

[22] F. Kelly, "Routing and capacity allocation in networks with trunk reservation," *Mathemat. Oper. Res.*, vol. 15, pp. 771–793, Nov. 1990.

[23] S.-P. Chung, A. Kashper, and K. Ross, "Computing approximate blocking probabilities for large loss networks with state-dependent routing," *IEEE/ACM Trans. Networking*, vol. 1, pp. 105–115, Feb. 1993.

[24] A. Acampora, "The multichannel multihop lightwave network," in *Proc. IEEE GLOBECOM'87*, vol. 3, Nov. 1987, pp. 1459–1467.

[25] R. K. Pankaj and R. G. Gallager, "Wavelength requirements of all-optical networks," *IEEE/ACM Trans. Networking*, vol. 3, pp. 269–280, Jun. 1995.

[26] V. Sharma and E. A. Varvarigos, "The ready-to-go virtual circuit protocol: A loss-free protocol for multigigabit networks using FIFO buffers," *IEEE/ACM Trans. Networking*, vol. 5, pp. 705–718, Oct. 1997.

[27] Y. Mei and C. Qiao, "Efficient distributed control protocols for WDM all-optical networks," in *Proc. Int. Conf. Computer Communication and Networks*, Sep. 1997, pp. 150–153.

[28] E. A. Varvarigos and V. Sharma, "An efficient reservation connection control protocol for gigabit networks," *Comput. Networks and ISDN Syst.*, vol. 30, pp. 1135–1156, Jul. 1998.

[29] J. Lang, "Almost-all optical switching: Architectures, protocols, and analyzes," Ph.D. dissertation, Univ. California, Santa Barbara, Dec. 1999.

[30] R. Koch, "Increasing the size of a network by a constant factor can increase performance by more than a constant factor," in *Symp. Foundations of Computer Science*, vol. 18, Oct. 1988, pp. 221–230.

[31] N. Antoniades, K. Bala, S. J. B. Yoo, and G. Ellinas, "A parametric wavelength interchanging cross-connect (wixc) architecture," *IEEE Photon. Technol. Lett,*, vol. 18, pp. 1382–1384, Oct. 1996.

Jonathan P. Lang (M'89) received the B.S. degree in electrical and computer engineering from the University of California at San Diego in 1993, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at Santa Barbara in 1995 and 1999, respectively.

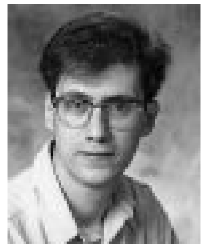
In 1999 he joined Calient Networks, Goleta, CA, as a Systems Engineer of Network Architecture, where he focuses on the design of Calient's network and routing architecture. Prior to joining Calient, he developed rate-allocation algorithms for ATM at Whitetree, Inc., Palo Alto, CA. As a Researcher and member of the MOST Center, University of California at Santa Barbara, he was responsible for developing the switch architecture and networking protocols for an all-optical packet switch. His research interests are in all-optical networking, protocol design, and optical switch architectures.

Dr. Lang is actively involved in the IETF and OIF.

Vishal Sharma (M'92) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1991, the M.S. degrees in signals and systems and computer engineering, both in 1993, and the Ph.D. degree in electrical and computer engineering in 1997, all from the University of California Santa Barbara.

From 1998 to 2000, he was with the Tellabs Research Center, Cambridge, MA, where he worked on the architecture and design of high-speed backbone routers, models for IP quality of service, and extensions to multi-protocol label switching for survivability and optical TDM networks. His current research interests include the design and analysis of high-speed communication networks, models for service differentiation in IP networks, and dynamic optical networks.

Dr. Sharma is a member of the IEEE Computer and Professional Communication Societies, and ACM Sigcomm. He is active in the CCAMP, MPLS, and Traffic Engineering Working Groups of the IETF.



Emmanouel A. Varvarigos (M'92) received the Diploma in electrical and computer engineering from the National Technical University of Athens, Athens, Greece, in 1988, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1990 and 1992, respectively.

During 1990, he was a Researcher with Bell Communications Research, Morristown, NJ. From 1992 to 1999, he was an Assistant and later an Associate Professor at the Department of Electrical and Computer Engineering,e University of California at Santa Barbara. During 1998–1999, he was a visiting Associate Professor at the Electrical Engineering Department, Delft University of Technology, the Netherlands. In 1999, he became a Professor in the Department of Computer Engineering and Informatics, University of Patras, Patras, Greece, and is currently the Director of the Hardware and Computer Architecture Division and the Head of the Communication Networks Lab. His research activities are in the areas of protocols and algorithms for high-speed networks, all-optical networks, high-performance switch architectures, parallel and distributed computing, interconnection networks, VLSI layout design, performance evaluation, and *ad-hoc* networks.

Dr. Varvarigos received a National Science Research Initiation Award and the First Prize in the National Competition in Mathematics.