Towards using the network as a switch: On the use of TDM in linear optical networks

David Karger [1] Muriel Médard [2]

[1] *karger@lcs.mit.edu*,

Computer Science and Artificial Intelligence Laboratory,

Massachusetts Institute of Technology

[2] *medard@mit.edu*,

Laboratory for Information And Decision Systems,

Massachusetts Institute of Technology

February 25, 2004

Abstract

A common problem in optical networking is that the large quantity of raw bandwidth available in such networks is often difficult to access. We show that time-division multiplexing (TDM) can be used to operate bus and ring architectures in a manner akin to a switch. We consider a time-varying approach, akin to that used in switching theory, instead of the static approach, related to the knapsack problem, often associated with fixed allocation of traffic onto circuits, often termed the grooming problem. Our approach is probabilistic in nature, and requires significant generalization beyond the Birkhoff-von Neumann statistical multiplexing approaches have that have been successful in switching theory. Our techniques rely on decompositions of fractional matchings (for architectures without erasures) and fractional interval graph colorings (for architectures with erasures) into integral matchings and colorings. We show, moreover, that such TDM using statistical multiplexing substantially reduces the amount of hardware (particularly ADMs) needed to utilize fully the available bandwidth in a range of optical networks. We show that a significant fraction (and in some cases all) of the bandwidth available to the system can be utilized, even if each node in the system has only a single ADM.

Keywords: mathematical programming, stochastic processes.

1 Introduction

Optical networks using wavelength division multiplexing (WDM) provide large capacity, but providing access to that capacity in in a way that is simple enough to be economical is challenging. In particular, the large quantity of aggregate capacity available in such networks is often inaccessible owing to constraints imposed by interactions among optical fibers, wavelengths on these fibers, electronic routers, and add-drop multiplexers (ADMs). This problem impacts both the design of optical networks (to make good use of their raw bandwidth) and the algorithms for fitting offered traffic into the network. For example, a particular placement of traffic on wavelengths that respects the (electrical) bandwidth at each given node may nevertheless be infeasible if the traffic destined for a node is spread out over more wavelengths than the node has wavelength ADMs.[1]

In this paper, we explore the possibility that reservation schemes implemented using time division multiplexing (TDM) using statistical multiplexing. We define to be *bandwidth feasible* any collection of traffic paths that (i) adds and drops no more traffic at a node than that node has bandwidth (this typically reflects the node's limited ability to handle traffic in the electrical domain) and (ii) places no more traffic on a fiber than the total bandwidth of the fiber. This definition ignores all consideration of wavelength conversion or dropping. We then explore the extent to which such bandwidth-feasible solutions can be implemented in systems that (i) do assign their traffic to distinct optical wavelengths and (ii) do so using only a limited number of ADMs per node. Broadly speaking, we seek to determine whether, for linear architectures such as buses and rings, TDM using appropriate statistical

[1] In this paper we consider an ADM to comprise single wavelength-tunable receiver/transmitter (transceiver) pair along with the attending equipment, such as filters, to access the fiber. As we discuss in Section 3.1, the transmitter or receiver may under certain conditions be constrained to a single wavelength without any loss in performance.

multiplexing can enhance our bandwidth utilization of the network, as well as reduce the amount of hardware required to access the bandwidth.

Birkhoff-von Neumann [Bir46], [vN53] statistical multiplexing using bipartite graph decompositions have been effective in switching applications [CCH00], [CLJ01], [LA], [LA01], [CCH]. However, as we shall show in this paper, the type of bipartite decompositions upon which Birkhoff-von Neumann techniques are based may not be sufficient in networks, and we require other decomposition approaches. Indeed, the topology of networks and issues of wavelength access may place constraints that are not present in switches, thus disallowing simple bipartite matches.

To address our question, we develop two families of techniques. The first family uses decomposition of weighted matchings into integral matchings. The second family of techniques uses interval graph coloring and randomized methods. While these two families of techniques have different lineages, the results they yield for applicable networks share a common theme. With proper scheduling, the use of only a few ADMs is sufficient to accommodate static traffic matrices that are feasible or close to feasible (in a way that we make precise later) with an arbitrarily large number of ADMs. Note that these results in effect imply that the issue of traffic grooming to reduce the required number of ADMs, important in static traffic assignment problems, is largely obviated in the case of dynamic scheduling.

We apply our techniques to some canonical linear architectures, which we detail in Section 2. In particular, we consider buses and rings. For buses, we consider both erasure-free models (where traffic on a wavelength remains on that wavelength until the wavelength is terminated at the end of the bus) and erasable models (where a node that is reading traffic from a wavelength has the option of erasing it to free that wavelength for use by other traffic downstream on the bus).

A sampling of our results is reflected in the following theorems. We consider architectures in which the bandwidth at each node is equal to a single wavelength; below we will discuss generalizations to larger bandwidths.

Theorem 1.1 *In a folded bus architecture without erasures, any bandwidth-feasible traffic pattern can be routed using a TDM scheme so long as each node has one ADM per wavelength's worth of traffic.*

This theorem holds for arbitrary node capacities so long as we give the node the same number of ADMs as are needed to carry its capacity needs. For example, if a node's capacity is equal to 2 wavelengths worth of traffic, it should be given 2 ADMs to make the conditions of the theorem hold.

Theorem 1.2 *In a folded bus architecture with erasures, given any bandwidth feasible traffic pattern, it is possible, using TDM and two ADMs per node, to route* $4/10$ *of the bandwidth of each connection. If the number of ADMs is increased to 4, it is possible to route* $7/10$ *of each connection. In a dual bus, the same can be achieved with twice as many ADMs.*

Theorem 1.3 *In a unidirectional ring architecture with erasures, given any bandwidth feasible traffic pattern, it is possible, using TDM and one ADM per node, to route* $1/5$ *of each connection. If the number of ADMs is increased to 2, it is possible to route* $2/5$ *of each connection.*

These theorems essentially state that, by giving up a small constant factor of the bandwidth of the network (generally a cheap resource), we can reduce the amount of hardware needed in the form of ADMs.

We first consider, in Section 3.1, to what extent a network using a reservation system implemented through TDM among integral matchings is effective in using the capacity of buses without erasures. All of our results hold as long as the total demand to or from a node is below the capacity of a wavelength, independently of the number of wavelengths. In the erasure-free model, we show that any routing pattern that is feasible with respect to aggregate capacity can be routed by a simple TDM scheme that alternates among several matchings of wavelengths to receivers and transmitters when two ADMs per node are used. In particular, we show that, using an appropriate reservation scheme, we can successfully serve any traffic matrix among nodes in the bus that can be accommodated using a multicommodity flow in which every node is a source and the only sink is the head end. Multicommodity flow determines the limitation on the network when the only constraint is the total available capacity on the bus, i.e. the total capacity of the wavelengths used by our WDM bus. Note that, when we use a TDM scheme as described above, any multicommodity flow is achievable when every node is equipped with a number of ADMs equal to the number of wavelengths. In effect, such a scenario removes issues of limited tunability at sender or transmitters.

In Section 3.2, we consider a family of techniques that combines interval graph coloring and randomized path selection for TDM. These techniques are applicable both to linear networks with and without erasures. In the erasable case, we show that a constant fraction of the traffic that is achievable using multicommodity flow can be accommodated by using TDM with one ADM per node per bus. Increasing the number of ADMs per node increases this fraction, but the incremental benefit of adding more ADMs decreases sharply.

Thus, if the traffic demand to or from each node is below the capacity of a wavelength, allowing a node to access more than two wavelengths at a time for reading or writing does not fundamentally allow for more efficient use of capacity in our system. We present our conclusions and discussion of extensions in Section 4.

2 Background and Model

2.1 Background

We consider a linear architecture in which access nodes write or read using tunable ADMs. Each ADM writes a single wavelength at a time and reads from a single wavelength at a time. The tunable ADM assumption means that any ADM can access any one wavelength for reading or writing at any time. Note that reading and writing are independent; thus the wavelengths a node accesses for reading or writing may be different. We allow reconfiguration of the ADMs and our consider a dynamic approach that makes use of such reconfiguration. We are interested both in maintaining a low number of ADMs at each access node and in accommodating data into the network. We also seek to characterize the relation and possible trade-off between these two goals.

The problem of reducing the amount of required hardware has been generally studied in the context of grooming in optical networks. Grooming is inherently a static problem, that addresses the problem of packing fractional traffic onto wavelengths and creating a routing wavelength assignment. Such a problem in relevant in cases where reconfiguration of nodes is done very infrequently. In its simplest form, the grooming problem is essentially a variant of the knapsack problem, in which elements (traffic) must be placed efficiently into the fewest possible knapsacks (wavelengths), according to some constraints. Reducing the number of wavelengths entails a reduction in the number of associated hardware, such as expensive ADMs. Extra constraints, such an unpacking as few knapsacks as possible (reading as few wavelengths as possible at another node), gradually increase the difficulty of the problem. The grooming problem inevitably inherits the complexity of the knapsack problem. The problem of grooming has generally been examined in the context of rings [DR02], [WCVM01], [GRS00], [MC98], [SGS99], [ZQ00] [CWM01], [BM00], [GSZF01], [Yoo01], [NKS99], [CL01], [LWL00], [WCF00] with some results being available for mesh networks [ML01], [KC01], [ZM02]. Note that, while the issue of grooming relies on variants of the knapsack problem, an inherently deterministic problem, the dynamic approach we take in this paper heavily relies on stochastic methods.

For static traffic matrices, the solutions are also generally static, with ADMs accessing fixed wavelengths and traffic being multiplexed in bandwidth. Our approach, in this paper, is instead to consider a rotation among a number of configurations, in each of which no traffic multiplexing is performed. Our results indicate that two ADMs per node will always achieve efficiency that is within a constant factor of the optimal solution that would be achieved with a number of ADMs equal to the number of wavelengths. Moreover, this results does not depend on the number of wavelengths or nodes. A similar approach has recently been proposed in [**?**]. That work also uses a scheduling approach to use the network, in effect, as a switch. Trees are used to perform routing in the network and the scheduling is performed by decomposing the traffic along such trees. The main differences between [**?**] and our approach are that we consider linear architectures, with no switches or other cross-connects in the interior of the network, since we address only linear networks, and we do not consider issues of delay. The latter issue becomes particularly significant if we consider wide-area networks. On the other hand, the linear architectures we consider are, as we discuss in Section II.B, mostly relevant to local and metropolitan area networks.

The families of techniques we use rely on matchings and on interval graph coloring. These techniques differ from the integer linear programs usually associated with traditional grooming problems. We should note that the type of graph coloring issues we consider are quite different from the type generally associated with solving the problem of wavelength assignment algorithms. In the latter context, a color is associated with a single wavelength and seek to prevent two paths on the same wavelength from occurring on the same link. Paths are thus created, generally through mesh networks, according to the wavelength continuity constraint and according to the fact that a single path occupies a whole bandwidth. The main concern in these problems is generally to reduce the number of wavelengths required to accommodate a certain traffic matrix. The type of coloring problems we address in this paper concern non-overlapping intervals that are directly linked to the linear architectures we consider. We create all possible allowable interval graph colorings, show expected behavior in terms of allowable traffic matrices, and then generate TDM schedules by selecting randomly from among the ensemble of allowable interval graph colorings. We do not seek to reduce the number of wavelengths used, but rather concern ourselves with attaining traffic matrices that are a significant fraction of those achievable when every node has an ADM for every wavelength.

2.2 Architectures for linear networks

In this Section, we present examples of the types of architectures for which our techniques apply. Our results apply to linear architectures, in which there is a clear linear ordering in the nodes of the network. Typical networks of this type are rings and buses, which we discuss below. In such networks, each node has exactly two contiguous nodes. The examples below are not intended to be exhaustive, but simply to provide a clear illustration of how ADMs are used in these architectures. We illustrate networks both with erasures and without erasures. In networks without erasures, the traffic destined to a node is read but not removed by that node from the traffic stream (traffic is flushed from the system at the end of the network to prevent recirculation of old traffic). In networks with erasures, a node removes from further circulation in the network the traffic that is intended for it. The techniques we develop later address both these types of networks.

We consider dual and folded bus architecture for optical access, as well as rings. Dual and folded bus architectures have been proposed for fiber-based local and metropolitan area networks such as distributed queue dual bus (DQDB) ([IEE], [Won89], [Bis90], [CGL90], [CGL91], [HM90], [vA90], [MB90], [Kam91], [Rod90], [WT93]), cyclic reservation multiple access (CRMA) ([Hua94], [SS94], [Nas90], [vALZZ91]), helical LAN (HLAN) ([Fin95], [RFB$^+$97], [BCH$^+$96]) and optical reservation multiple access (ORMA) ([Ham97]). Recently, bus schemes that are resilient to a single link or node failure have been proposed in [MLL02], [LM01], which also give an overview of optical bus schemes. An advantage of optical bus schemes is that they they do not require nodes to perform any switching or routing. Access nodes need only place traffic onto the bus and read it off the bus. Rings, both unidirectional and bidirectional, are well known in the literature [Wu92].

In the non-erasure model, the nodes do not erase traffic when they read from the bus. Instead, all read traffic is expunged at the end of the bus. The constraint that nodes not erase traffic is related to hardware simplicity. Erasing traffic generally requires light-terminating equipment that transfers all traffic from the optical traffic to the electronic domain. In the electronic domain, certain portions of the traffic are removed and then a new optical traffic stream is generated. Erasing traffic while remaining all-optical is possible ([OOB00], [BOR$^+$00]), but is generally onerous. Reading can be done using a passive optical tap that does not affect the optical stream except for a slight degradation in signal strength.

We have n access nodes, each with read and write capability over two wavelengths simultaneously. We have m wavelengths, that constitute the folded bus. Each wavelength has capacity C, say in bits per second. We consider that we have a traffic demand matrix D whose elements $d_{i,j}$ are the demand of traffic demand from node i to node j. Note that the diagonal of D has zero elements. The demand is expressed also in terms of desired capacity in the same units as C.

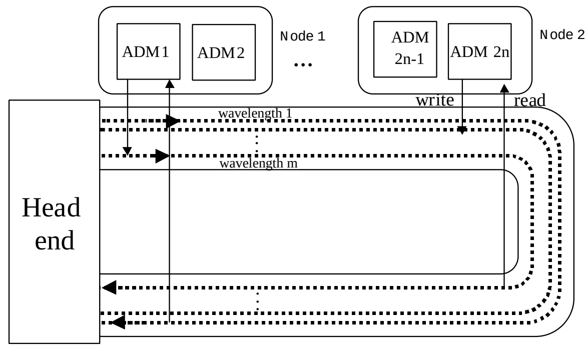
We present both folded and dual bus topologies without erasure. We consider that every node is equipped with two ADMs *per bus*. Thus, in the case of a folded bus, each node has two ADMs per node. In the case of a dual bus, we consider that we have two buses and each node presides over four ADMS.

A folded bus emanates from a single head-end and collects traffic by traversing all the nodes in the bus. After collecting traffic from all the nodes, the folded bus reverses directions and distributes the traffic to all the nodes in the reverse order from which it collected. After distributing to all the nodes, the bus terminates, purging all of its traffic. Figure 2.2 shows our folded bus under two different reading and writing configurations. We consider systems with and without erasures, as detailed below. We show a single ADM per node active. Each ADM writes on the top part of the bus and read on the bottom part of the bus.

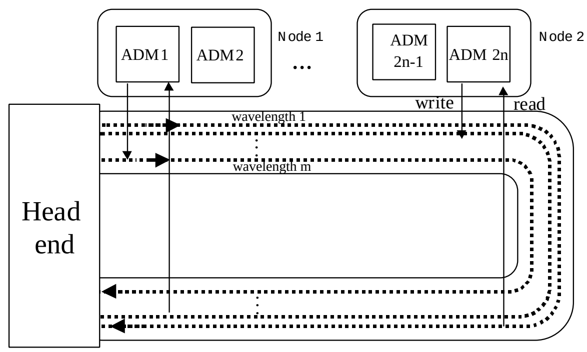
A dual bus consists of two unidirectional buses between two head ends. Figure 2.2 shows a single node with four ADMs. Only two of the ADMs are shown as active. ADMs 1 and 2 access the top bus for writing and the bottom bus for reading. ADMs 3 and 4 access the bottom bus for writing and the top bus for reading.

If erasures are permitted, then nodes can erase traffic intended for them and that capacity can be used by downstream nodes. We model ADMs in this case as writing to one wavelength and reading, with erasure, to a single wavelength. The capacity freed by erasures is available to downstream nodes. Note that we may extend our model to other architectures, such a buses where we use a single transmitter and two receivers, which requires fewer components than two full ADMs. The dual bus for the case with erasures resembles in configuration the bus shown in Figure 2.2, except that erasures are now possible.

Unlike buses, rings inherently require erasures, since they do not have a natural starting or ending point. Nodes terminate and erase the traffic that is destined to them. We do not show a ring, since the topologies of unidirectional and bidirectional rings are well known.



a. Configuration in which ADM 1 in node 1 writes to wavelength m and reads from wavelength 1,
ADM 2n in node n writes to wavelength 2 and reads from wavelength m



b. Configuration in which ADM 1 in node 1 writes to wavelength m and reads from wavelength 2,
ADM 2n in node n writes to wavelength 2 and reads from wavelength 1

Figure 1: Folded bus with two different read and write configurations. The wavelengths 1 through m are indicated as thick dotted lines.

3 Effectiveness of TDM reservations schemes

3.1 Bus without erasures

In this Section, we show that the fact that an access node can only read and write to a single wavelength at a time does not affect how efficiently we use the available capacity. We consider a reservation mode, in which nodes alternate among predetermined configurations, as shown in figure 2.2. Since nodes do not erase traffic, their relative positions do not affect possible traffic assignments. When we consider the effect of erasures, the relative placement of nodes is important. We show that any traffic matrix D can be accommodated arbitrarily well by alternating among predetermined assignments. Each assignment can be viewed as a mapping from transmitter

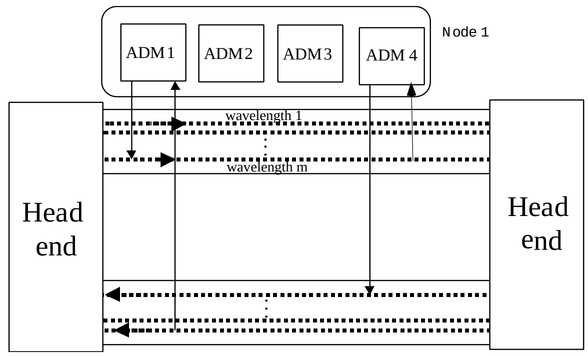


Figure 2: Dual bus in one read and write configuration. The wavelengths 1 through m are indicated as thick dotted lines.

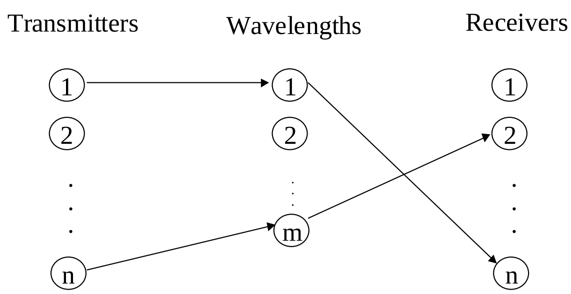


Figure 3: Matching of transmitter → wavelength → receiver. In this example, node 1 transmits to node n on wavelength 1 and node n transmits to node 2 along wavelength m.

nodes to receiver nodes via wavelengths, as shown in figure 3.1. Thus, the assignment of wavelengths for reading and writing is a matching transmitter → wavelength → receiver of the type shown in figure 3.1. At any one time, a node writes to a whole wavelength and reads a whole wavelength. Thus, a single arc may emerge from any transmitter, a single arc may enter any wavelength, a single arc may emanate from any wavelength and a single arc may enter any receiver.

In the case of a bus without erasures, a bandwidth-feasible demand is easy to describe. It essentially consists in removing the constraints of access to specific wavelengths. A bandwidth-feasible is simply a collection of pairwise demands such that (i) the demand being sent by any node and the demand being received by any node is at most one and (ii) the total traffic being routed is at most m. We can represent this demand matrix as a bipartite graph with one side S containing one node per sender, the other side R containing one node per receiver, and and edge from $s \in S$ to $r \in R$ receiving weight equal to the demand (in units of a wavelength) from s to r. It follows from bandwidth feasibility that the total weight incident on any vertex $s \in S$ or $r \in R$, which we call the *fractional degree* of the vertex, is at most 1, while the total weight of edges is at most m. On this bipartite graph, an "instantaneously feasible routing" that represents who is sending and receiving what at any instant, is also easy to describe: we draw an edge from each sender in S to the corresponding receiver in R.

Let us first consider the case where each node has but one ADM. Since each node can receive and transmit only one wavelength, the set of edges drawn will form a *matching*, and since the total number of wavelengths is at most m, the *size* of the matching (number of edges) will be at most m.

A TDM of instantaneously feasible routings corresponds to an assignment to each possible matching M a time slice $t_M \leq 1$ representing how much of the unit interval is spent routing matching M. To satisfy the demand d_{sr} from sender s to receiver r, it is necessary that

$$\sum_{M \ni (s,r)} t_M \geq d_{sr}.$$

To fit in the interval it is of course necessary that $\sum_M t_M \leq 1$. Since it doesn't hurt to route extra traffic, we can always add arbitrary additional matchings and thus assume that $\sum_M t_M = 1$. An assignment of such values t_M is known as a *convex combination of matchings.*

The question of whether it is possible to route a particular demand matrix can be rephrased as asking whether there is a convex combination $\{t_M\}$ of m-edge matchings such that $\sum_{M \ni (s,r)} t_M \geq d_{sr}$ for every pair (s, r).

We begin by proving an easy case. Suppose that there are exactly m senders and m receivers, and that each sender is sending (and each receiver is receiving) exactly one unit. This implies that the total demand being sent is m, so the demand matrix is feasible. Note, however, that possibly all m^2 demands might be nonzero, so the traffic cannot necessarily be routed via a single fixed assignment of wavelengths to demands. Nevertheless, there is always a feasible TDM solution in this case. This follows from a well known theorem: any bipartite graph in which all fractional degrees are exactly 1 can be written as a convex combination of matchings.

This fact is often stated in slightly different terms. If we view the bipartite graph as an $m \times m$ *adjacency matrix* in which the (s, r) entry represents the demand from s to r, then the unit-degree requirement says that the matrix is a so-called *stochastic matrix* (all row and column sums are one). It is a well-known theorem, often referred to as Birkhoff-von Neumann, that any such stochastic matrix can be represented as a convex combination of *permutation matrices* (exactly one 1 in each row and column) ([Bir46], [vN53]). Such a permutation matrix is the adjacency matrix for a matching.

The claimed decomposition into matchings can be found in polynomial time via bipartite matching algorithms; it assigns nonzero weight to at most m^2 distinct matchings.

Since there are m senders and receivers, each perfect matching has m edges and is therefore feasible as an instantaneous routing. This shows that a TDM solution exists.

It is interesting to note that, since it is irrelevant which pair is using which wavelength, we can choose the colors in each TDM slice such that the i^{th} sender is always using color i. It follows that only the receivers need to be tunable in this approach. Alternatively, only the transmitters need to be tunable.

Now let us consider the general case in which there are more than m senders and/or receivers. As above, our goal is to write the demand graph as a convex combination of m-edge matchings. We begin by adding artificial demand to the graph to increase the degree of every vertex to one (note this can be done since, so long as one side has a vertex of degree less than one, the other will too). The resulting stochastic matrix can be expressed as a convex combination of perfect matchings in the augmented graph. If we remove the artificial edges from each matching, we find that our graph can be expressed as a convex combination of (not necessarily perfect) matchings.

Since the demand graph has total weight m, it follows that the average number (under the same convex combination) of edges in the matchings is m. This is almost what we want, but we cannot settle for an average. Suppose it has more (at least $m+1$). Then by average, some other matching must have less than m (at most $m-1$). We will show how to move some edges from the larger to the smaller matching, decreasing the gap between large and small matchings. This process can be repeated until all matchings have the same size (which by the averaging analysis must be m).

So consider the union of these two matchings. The maximum degree (number of edges) of any vertex in the union is at most 2, so the union consists of a collection of disjoint cycles and paths. We go through each cycle and path in turn, assigning edges alternately to two sets. The fact that each set is a matching follows immediately from the fact that we are taking alternate edges from (disjoint) paths and cycles. If the total number of edges is even, both matchings end up with the same number; if odd, they are off by one. In any case, we can continue this redivision until no two matchings differ in size by more than one edge; as was argued above, this implies that all matchings have size m.

From the fact that there is a convex combination of matchings of size exactly m, it is immediate from standard polyhedral analysis that the combination can be made up of a number of distinct matchings no greater than the number of demand edges plus the number of nodes, which in turn is at most twice the number of demand edges.

The above approach generalizes to the case where nodes have bandwidth to handle more than one wavelength at a time. If a node has k ADMs (and the bandwidth to handle them all) we can simply treat it as a set of k independent nodes each with one ADM and one unit of bandwidth.

3.2 Bus with erasures

We now consider the case where nodes can erase traffic after they read it. We introduce an interval-graph coloring technique to route traffic from fractionally feasible solutions. Our results are weaker than the erasure case: we do not prove that with a small number of ADMs, any fractionally feasible demand matrix can be routed. But we show that the gap is limited: at least a constant fraction (independent of m) of each demand from the fractional pattern can be routed is TDM.

We begin by describing a "bandwidth-feasible" solution. We focus on the case where the transmitters and receivers of the ADM are laid out linearly on a fiber bus (we will discuss rings in a later section). Given such a layout, there is exactly one path between a given transmitter and receiver, which must consume bandwidth on the portion of the fiber running from the transmitter to the receiver. Thus, each demand between a pair of ADMs imposes a load on the fibers between the ADMs. We can now state the obvious bandwidth-feasibility restrictions: (i) the total demand incident on any transmitter or receiver it at most 1 and (ii) the total demand that traverses any particular point on the fiber is at most m.

To determine a routing for such a bandwidth-feasible demand, we represent the demand as an interval graph. We number the transmitters and receivers in the order they appear on the bus. The demand between i and j is represented as the interval (i, j) with weight equal to the demand. If two intervals overlap, their demands cannot simultaneously be transmitted on the same wavelength. On the other hand, two intervals that do not overlap *can* have their demand transmitted simultaneously on the same wavelength thanks to the erasures. Let us call a set of mutually disjoint intervals (which can all simultaneously use the same wavelength) a *fragment*.

As observed above, the maximum total weight of intervals crossing any particular point of the fiber is at most m. It is a well known result about interval graphs (following from the fact that such graph are *perfect*) that this implies that the set of intervals have *fractional chromatic number* at most m ([GS88], [GLL82]). This means the following: we can assign a weight w_F to each fragment F such that

- The total weight of fragments that contain interval e is equal to the weight of interval e (demand between the endpoints of e)
- The total weight of all fragments is m.

This fractional coloring can be computed by a simple quadratic-time greedy algorithm. It assigns nonzero weight to a number of fragments no greater than the number of nonzero demand intervals.

We can use the resulting coloring to define a TDM scheme. Since the total weight of all fragments is m, we can define a probability distribution over fragments in which the probability fragment F is chosen is w_F/m. Under this probability distribution, the probability that a fragment is chosen that contains a particular interval (demand) of weight w is simply w/m.

Consider choosing m fragments at random (with replacement) from this distribution. Among the intervals of the fragment, we will keep any interval whose endpoints are not being used by any other interval in the chosen fragments. The resulting set of intervals uses each node (ADM) at most once. Also, since they come from m fragments, they can all be simultaneously routed on m wavelengths (we use one wavelength for the surviving intervals from each fragment). Let us now analyze the probability that a particular interval I of weight w is routed. This will happen if and only if exactly one of the m selections chooses a fragment that contains I, while none of the other $m-1$ selections include any interval with the same endpoints as I. The probability that a particular choice contains I is precisely w/m. Since the total weight of intervals incident on each endpoint of I is at most 1, we deduce that the probability that a random fragment shares an endpoint with I is only $2/m$. There are of course exactly m possibilities for which of the m choices contains the interval I. Thus, we conclude that the probability that one selection contains I and the remainder do not contain either endpoints of I is precisely

$$m(w/m)(1-2/m)^{m-1} \approx we^{-2}$$

Now let us imagine performing some large number N of these "m-choice" experiments, and setting up a TDM in which each resulting routing persists for $1/N$ time units (so we get a TDM over the span of one time unit). By the above analysis, the expected number of these experiments that routes the interval I is Nwe^{-2}. By the Chernoff bound, as N grows, the actual number of experiments that routes I will be arbitrarily close to Nwe^{-2}. It follows that the total time spent routing I during the time unit will be we^{-2}. Since the demanded time is w, it follows that we are routing ane^{-2} fraction of the demand of interval I. Rephrasing, we see that we can route an e^{-2} fraction of all demands of any bandwidth-feasible traffic matrix.

We can strengthen this result with a more careful algorithm. Instead of discarding any interval whose endpoints are shared with another chosen interval, we note that our m fragments are chosen in order, and keep an interval if none of the *previously chosen* fragments uses its endpoints. It is easy to see that this produces a jointly feasible set of intervals for simultaneous routing on m wavelengths. The probability that interval I is routed in the $(k+1)^{st}$ fragment of this scheme is then equal to the probability that the prior k fragments did not contain an interval using the endpoints of I, which is at most $(1-2/m)^k$, times the probability that a fragment containing I is chosen, which is w/m. Summing over possible k, the overall probability that I is chosen is

$$\begin{aligned} & \sum_{k=0}^{m-1}(1-2/m)^k(w/m) \\ = \quad & (w/m)\frac{1-(1-2/m)^m}{2/m} \\ \approx \quad & \frac{w}{2}(1-e^{-2}) \\ \approx 0.43 \cdot w & \end{aligned}$$

In other words we can route roughly 4/10 of any bandwidth feasible routing.

Instead of scaling a bandwidth-feasible routing, one might prefer to ask directly *which* bandwidth-feasible routings can actually be routed. We consider two parameters of the routing: (i) the maximum total bandwidth $m' \leq m$ induced on any portion of the fiber and (ii) the maximum bandwidth $b < 1$ incident on any node. We can form an interval graph for this instance, give it an m'-fractional coloring, and from this derive a probability distribution in which an interval of weight w is selected with probability w/m'. Applying the analysis above, we find that the probability of routing a particular interval when we select m fragments is

$$\sum_{k=0}^{m-1}(1-2b/m')^k(w/m') \quad = \quad (w/m')\frac{1-(1-2b/m')^m}{2b/m'}$$

$$\approx \frac{w}{2b}(1 - e^{-2bm/m'})$$

which exceeds w whenever

$$\begin{aligned} 1 - e^{-2bm/m'} &> 2b \\ \frac{m}{m'} &> \frac{1}{2b} \ln \frac{1}{1-2b} \end{aligned}$$

Thus, for example, we can route any demand matrix for which $b = 1/3$ and $m' \leq .6m$.

We can extend the above result by supposing that each node has more than one ADM. In this case, as we select random fragments, we can keep an interval so long as each of its endpoints still has an ADM to spare. For example, if we allow 2 ADMs per node, we can keep I so long as at most one preceding interval uses the endpoints of I. This probability can be upper bounded after a lengthy calculation by $(1 - 2e^{-2}) \approx .73$ times the weight of the interval, meaning that we can route roughly 7/10 of any bandwidth feasible demand using 2 ADMs.

3.3 Rings

Rings have been the subject of much study in the area of grooming, which is related to our study insofar as it seeks to accommodate traffic onto wavelengths while reducing the cost of hardware, often indicated by the number of ADMs required. As discussed in Section 2, the approach we take is significantly different from the grooming one. Note that our use of decomposition into intervals is markedly different than that presented in [DR02]. In the latter work, the decomposition is used to approximate an ILP to determine the most benefit that may be derived from routing. In our work, it is not a means of creating a schedule among several integral solutions.

Our interval graph coloring technique generalizes quite straightforwardly to rings. Consider a unidirectional ring (eg, one in which all traffic is routed clockwise). There is a unique path between any sender and receiver in such a ring, and it corresponds to an arc on the circle. In such an architecture, erasures, or course, are a must. As we did with interval graphs for the bus, we can define a *circular arc graph* that captures which demand pairs can be simultaneously transmitted on the same wavelength ([GLL82]). We can apply the same technique of coloring the demand arcs to produce fragments, and choosing a collection of m fragments to route simultaneously.

A small problem arises, however, in that (fractional) coloring on circular arc graphs is more complex than on interval graphs. It is no longer the case that the fractional chromatic number is equal to the maximum weight of all arcs traversing a certain point of the circle. In fact, finding the optimal coloring of a circular arc graph is NP-complete. However, we can recover by slightly weakening our results. Consider deleting a particular point on the ring, along with all the demands that cross it. What is left is an interval graph, and our previous analysis applies to it. On the other hand, the set of deleted demands, since they all share a point on the fiber, have total weight at most m, and so can be fractionally colored with m colors. This shows that any circular arc graph with maximum demand m at any point can be fractionally colored with $2m$ colors. This is enough to let us apply our previous techniques, sacrificing a factor of 2 in the results since we have $2m$ rather than m colors. A more sophisticated analysis of the gap between the maximum demand at a point and the fractional chromatic number lets us replace the factor 2 by 3/2 [Kar80] .

Our results apply equally to a bidirectional ring since we can simply analyze each direction of the ring independently.

4 Conclusions

We have investigated, for a variety of WDM network linear topologies, such as buses and rings with and without erasures, the efficiency of reservation schemes implemented using TDM. We have shown that equipping every node with a number of ADMs equal to the number of wavelengths required to carry the traffic to or from a node can accommodate traffic matrices that are of the same order of those achievable with an unlimited number of ADMS for non-erasure models. For erasure models, this number of ADMs can accommodate a fixed fraction of the traffic achievable under the unlimited ADM scenario.

Our results, surprisingly, hold regardless of the number of nodes and wavelengths on the network. This independence differs markedly from the relation that appears in traditional grooming between, on the one hand, the number of ADMs and, on the other hand, the number of nodes and wavelengths. Thus, not only is our approach markedly different from traditional grooming, but the results we obtain are also very different.

The other surprising result is that we do not require any intermediate nodes to perform any actions. In particular, it is not necessary for intermediate ADMs to act as wavelength changers. All traffic between two nodes at any time is thus handled by a single uninterrupted lightpath.

Our work indicates that the use of time-division multiplexing may allow a reduction in the number of ADMs that are used in networks. In particular, a moderate number of ADMs may prove, for linear networks, to be optimal or within a significant factor of optimal. While the methods we have presented may not in general extend directly to arbitrary mesh networks, the essence of the TDM approach may possibly allow operation with a small number of ADMs.

References

[BCH+96] R.A. Barry, V.W.S. Chan, K.L. Hall, E.S. Kintzer, J.D. Moores, K.A. Rauschenbach, E.A. Swanson, L.E. Adams, C.R. Doerr, S.G. Finn H.A. Haus, E.P. Ippen, W.S. Wong, and M. Haner. All-optical network consortium-ultrafast TDM networks. *IEEE Journal on Selected Areas in Communications*, 14(5):999 –1013, June 1996.

[Bir46] G. Birkhoff. *Three Observations on Linear Algebra*. Universidad Nacional Tucumán, Series A 5, 1946.

[Bis90] C. Bisdikian. Waiting time analysis in a single buffer DQDB (802.6) network. In *Proceedings IEEE INFOCOM*, pages 610–616, 1990.

[BM00] R. Berry and E. Modiano. Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic. In *IEEE Journal on Selected Areas in Communications*, October 2000.

[BOR+00] D.J. Blumenthal, B.-E. Olsson, G. Rossi, T.E. Dimmick, L. Rau, M. Masanovic, O. Lavrova, R. Doshi, O. Jerphagnon, J.E. Bowers, V. Kaman, L.A. Coldren, and J. Barton. All-optical label swapping networks and technologies. *IEEE Journal of Lightwave Communications*, 18(12):2058 –2075, December 2000.

[CCH] C.-S. Chang, W.J. Chen, and H.-Y. Huang. Birkhoff-von neumann input-buffered crossbar switches for guaranteed-rate services. *IEEE Transactions on Communications*.

[CCH00] C.-S. Chang, W.J. Chen, and H.-Y. Huang. Birkhoff-von neumann input buffered crossbar switches. In *Proceedings IEEE INFOCOM*, volume 3, pages 1614–1623, 2000.

[CGL90] M. Conti, E. Gregori, and L. Lenzini. DQDB under heavy load: performance evaluation and fairness analysis. In *Proceedings IEEE INFOCOM*, pages 133–145, 1990.

[CGL91] M. Conti, E. Gregori, and L. Lenzini. A methodological approach to an extensive analysis of DQDB performance and fairness. *IEEE Journal on Selected Areas in Communications*, 9(1):76–87, January 1991.

[CL01] T.Y. Chow and P. Lin. The ring grooming problem. In *submitted to SIAM Journal of Discrete Mathematics*, volume 2001.

[CLJ01] C.-S. Chang, D.-S. Lee, and Y.-S. Jou. Load balanced birkhoff-von neumann switches. In *IEEE Workshop on High-Performance Switching and Routing*, pages 276–280, 2001.

[CWM01] W. Cho, J. Wang, and B. Mukherjee. Improved approaches for cost-effective grooming in WDM ring networks: uniform traffic case. *Photonic Networks and Communications*, 3(2):245–254, April 2001.

[DR02] R. Dutta and G.N. Rouskas. On optimal traffic grooming in WDM rings. *IEEE Journal on Selected Areas in Communications*, 20(1):110–121, January 2002.

[Fin95] S.G. Finn. HLAN - an architecture for optical multi-access networks. In *Digest of the LEOS Summer Topical Meetings*, pages 45 –46, 1995.

[GLL82] U.I. Gupta, D.T. Lee, and Y. Y.-T. Leung. Efficient algorithms for interval graphs and circular-arc graphs. In *Networks*, volume 12, pages 459–467, 1982.

[GRS98] O. Gerstel, R. Ramaswami, and G. Sasaki. Cost effective traffic grooming in WDM rings. In *infocom*, pages 94–101, March 1998.

[GRS00] O. Gerstel, R. Ramaswami, and G. Sasaki. Cost effective traffic grooming in WDM rings. *IEEE/ACM Transactions on Networking*, 8(5):618–630, October 2000.

[GS88] M. Grötschel and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer Verlag, Berlin, 1988.

[GSZF01] H. Ghafouri-Shiraz, G. Zhu, and Y. Fei. Effective wavelength assignment algorithms for optimizing design in SONET/WDM rings. *IEEE Journal of Lightwave Communications*, 19(10):1427–1439, October 2001.

[Ham97] M. Hamdi. ORMA: a high-performance MAC protocol for fiber-optic LANs/MANs. *IEEE Communications Magazine*, 35(3):110–119, March 1997.

[HM90] E.Y. Huang and L.F. Merakos. On the access fairness of the DQDB MAN protocol. In *Proceedings of IPCC*, pages 325–329, 1990.

[Hua94] E.Y. Huang. Analysis of cyclic reservation multiple access protocol. In *19th Conference on Local Computer Networks*, pages 102–107, 1994.

[IEE] Distributed queue dual bus (DQDB) - subnetwork of a metropolitan are network (MAN). In *IEEE Standard 820.6*.

[Kam91] A.E. Kamal. Efficient multi-segment message transmission with slot reuse on DQDB. In *Proceedings IEEE INFOCOM*, pages 869–878, 1991.

[Kar80] I. A. Karapetian. On coloring of arc graphs. *Dokladi (Reports) of the Academy of Science of the Armenian Soviet Socialist Republic*, 70(5):306–311, 1980.

[KC01] V.R. Konda and T.Y. Chow. Algorithm for traffic grooming in opticaql networks to minimize the number of transceivers. In *IEEE Workshop on High Performance Routing and Routing*, volume pages 218–221, 2001.

[LA] J. Li and N. Ansari. Enhanced birkhoff-von neumann decomposition algorithm for input queued switches. *IEE Proceedings on Communications*.

[LA01] J. Li and N. Ansari. Qos guaranteed input scheduled algorithms with low delay. In *IEEE Workshop on High-Performance Switching and Routing*, pages 412–414, 2001.

[LM01] A. Libarikian and M. Médard. Robustness of bus overlays in optical networks. In *Thirty-Ninth Annual Allerton Conference on Communication, Control, and Computing.*, 2001.

[LWL00] X.Y. Li, P.J. Wan, and L. Liu. Select line speeds for single-hub SONET/WDM ring networks. In *icc*, volume 1, pages 495–499, 2000.

[MB90] B. Mukherjee and S. Banerjee. Incorporating continutaion-of-message (COM) information, slot reuse, and fairness in DQDB networks. In *Division of Computer Science, University of California, Davis, CA, Technical Report CSE-90-42*, October 1990.

[MC98] E.H. Modiano and A.L. Chiu. Traffic grooming algorithms for minimizing electronic multiplexing costs in unidirectional SONET/WDM ring networks. In *Proceedings of Conference on Information Sciences and Systems*, March 1998.

[ML01] E. Modiano and P. Lin. Traffic grooming in WDM networks. In *IEEE Communications Magazine*, volume 39, pages 124–129, July 2001.

[MLL02] M. Médard, S.S. Lumetta, and L. Li. A network management architecture for robust packet routing in mesh optical access networks. *IEEE Journal of Lightwave Communications*, 20(4):822–833, May 2002.

[Nas90] M.M. Nassehi. CRMA:an access scheme for high-speed LANs and MANs. In *Proceedings of the IEEE International Conference on Communications*, volume 4, pages 1697 –1702, 1990.

[NKS99] T. Nakata, T. Kanetake, and S. Sasaki. Cost-merit estimation of introducing optical layer with efficient traffic grooming. In *ofc*, volume 2, pages 65–67, 1999.

[OOB00] P. Ohlen, B.-E. Olsson, and D.J. Blumenthal. All-optical header erasure and penalty-free rewriting in a fiber-based high-speed wavelength converter. *IEEE Photonics Technology Letters*, 12(6):663–665, June 2000.

[RFB+97] K. Rauschenbach, S. Finn, R. Barry, K. Hall, J. Moores, and N. Patel. 100-Gbit/s time-division multiplexed multi-access networks. In *Proccedings of OFC*, pages 86 – 87, 1997.

[Rod90] M.A. Rodrigues. Erasure nodes: performance improvements for the IEEE 802.6 MAN. In *Proceedings IEEE INFOCOM*, pages 636–643, 1990.

[SGS99] J.M. Simmons, E.L. Goldstein, and A.M. Saleh. Quantifying the benefit of wavelength add/drop in WDM with distance-independent and dependent traffic. In *jlt*, volume 17, pages 48–57, January 1999.

[SS94] O. Sharon and A. Segall. Schemes for slot reuse in CRMA. *IEEE/ACM Transactions on Networking*, 2(3):269–278, June 1994.

[vA90] H.R. van As. Performance evaluation of bandwidth balancing in the DQDB MAC protocol. In *Proceedings of EFOC/LAN Conference*, 1990.

[vALZZ91] H.R. van As, W.W. Lemppenau, P. Zafiropulo, and E.A. Zurfluh. CRMA-II: a Gbit/s MAC protocol for ring and bus networks with immediate access capability. In *Proceedings of the EFOC/LAN Conference*, pages 262–272, 1991.

[vN53] J. von Neumann. *A certain zero-sum two-person game equivalent to the optimal assignment problem*, volume 2 of *Contribution to the theory of games*. Princeton University Press, 1953.

[WCF00] P.J. Wan, G. Calinescu, and O. Frieder. Grooming of arbitrary traffic in SONET/WDM BLSRs. *IEEE Journal on Selected Areas in Communications*, 18(10):1995–2003, October 2000.

[WCVM01] J. Wang, W. Cho, V.R. Vemuri, and B. Mukherjee. Improved approaches for cost-effective grooming in WDM ring networks: ILP formulations and single-hop and multihop connections. *IEEE Journal of Lightwave Communications*, 19(11):1645–1655, November 2001.

[Won89] J.W. Wong. Throughput of DQDB networks under heavy load. In *Proceedings of EFOC/LAN Conference*, pages 146–151, 1989.

[WT93] G.C. Watson and S. Tohme. S++- a new MAC protocol for Gb/slocal area networks. *IEEE Journal on Selected Areas in Communications*, 11(4):531–539, may 1993.

[Wu92] T.-H. Wu. *Fiber Network Service Survivability.* Artech House, 1992.

[Yoo01] M.G. Yoon. Traffic grooming and light-path routing in WDM ring networks with hop-count constraint. In *icc*, volume 3, pages 731–737, 2001.

[ZM02] K. Zhu and B. Mukherjee. 2 traffic grooming in an optical WDM mesh network. *IEEE Journal on Selected Areas in Communications*, 20(1):122–133, January 2002.

[ZQ00] X. Zhang and C. Qiao. An effective and comprehensive approach for traffic grooming and wavelength assignment in SONET/WDM rings. *IEEE/ACM Transactions on Networking*, 8(5):608–617, October 2000.