Designing Minimum Cost Nonblocking Communication Networks

J. Andrew Fingerhut Subhash Suri Jonathan S. Turner

WUCS-96-06

February 5, 1996

Department of Computer Science
Campus Box 1045
Washington University
One Brookings Drive
St. Louis, MO 63130-4899

Abstract

This paper addresses the problem of topological design of ATM (and similar) communication networks. We formulate the problem from a worst-case point of view, seeking network designs that, subject to specified raffic constraints, are nonblocking for point-to-point and multicast virtual circuits. Within this model we give various conditions under which *star networks* are optimal or near-optimal. These conditions are approximately satisfied in many common situations making the results of practical significance. An important consequence of these results is that, where they apply, there is no added cost for nonblocking multicast communication, relative to networks that are nonblocking for point-to-point traffic only.

Designing Minimum Cost Nonblocking Communication Networks

J. Andrew Fingerhut Subhash Suri Jonathan S. Turner

1. Introduction

As computer networks get larger and more complex, the need for careful planning in the design and configuration stage becomes more and more important. This has become an issue even for conventional shared-access LAN and router based networks, but is even more crucial for ATM. Because ATM networks support virtual circuit routing and must provide quality-of-service guarantees to real-time traffic (voice, video, etc.), connection requests can block if the network backbone does not have sufficient available bandwidth to satisfy a user's needs.

The network design problem is not a new one. References [7, 8, 9, 13] are representative of the recent published research in this area. However, ATM networks differ from telephone and classical data networks in several ways. First, they are multirate networks, meaning that their virtual circuits can operate at any bandwidth from a few bits per second to over one hundred megabits per second. They can support a wide range of different applications with different bandwidth needs, different connection request rates and different holding times. Moreover, unlike traditional data networks they must be capable of providing connections with a guaranteed quality of service, requiring allocation of bandwidth to individual virtual circuits and raising the possibility of virtual circuit blocking. Second, ATM networks support not only point-to-point virtual circuits but also multicast. Multicast virtual circuits are essential for applications like video distribution or multimedia conferencing and can include both one-to-many and many-to-many transmission patterns. Finally, ATM networks are much less predictable than telephone networks or traditional low speed data networks. There is no reliable statistical data on application characteristics and connection request patterns. Indeed, the flexibility which is ATM's greatest strength makes it highly unpredictable, so classical network planning techniques which rely heavily on statistical analysis become less relevant. In ATM networks, the whole notion of blocking probability for virtual circuit setup must be called into question, since there is no reasonable possibility of validating the probabilistic assumptions that must go into any analysis of blocking probability.

Our model is inspired by and builds upon the classical theory of nonblocking switching networks developed by Beneš [1], Clos [3] and Pippenger [14], among others, and generalized

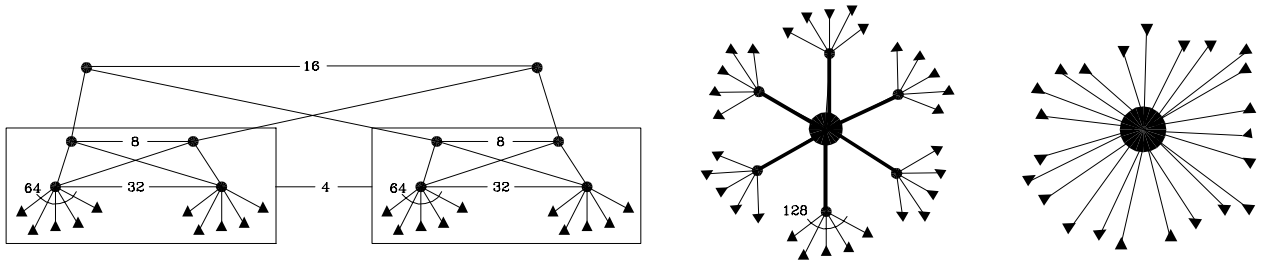

Figure 1: Alternative Campus Network Designs

to multirate switching networks by Melen and Turner [11, 12]. The present paper differs from the work in switching networks in that it addresses the design of networks with irregular topologies and traffic characteristics, and it takes into account the costs of transmission links spanning substantial geographical distances. It also differs from prior work in topological design of networks in allowing a much less constrained and detailed specification of traffic requirements [7, 9, 13]. One can think of our model as implicitly allowing the specification of a very large number of traffic matrices.

Before getting into the abstract formalization of the network design problem, it's useful to understand how different elements of a network contribute to its cost and what this can mean in the context of a specific instance of the network design problem. In ATM networks, there are three basic things that contribute to the cost of a system (1) fiber plant, (2) transmission electronics and (3) switching systems.

- *Fiber plant.* The cost of a pair of fibers in a typical multi-fiber cable is about \$0.30 per linear meter of fiber.[1] (the cost per pair is about 50% higher when purchased as a single pair cable, rather than in a large multi-fiber cable). This means that short lengths of fiber are inexpensive relative to other components of a system (\$75 for 250 meters, for example), but long lengths are relatively expensive (\$60,000 for 200 km). This elementary observation means that the optimal network topologies are qualitatively very different depending on the geographic distances that must be spanned.

- *Transmission electronics.* The transmission electronics includes the opto-electronic conversion circuits and the circuits that format the signal for transmission over a fiber and perform clock recovery and synchronization at the receiver. For moderate distances (say up to 10 km) the costs for these components is roughly \$500 per end for 150 Mb/s links. Higher speed interfaces currently have limited availability, but when widely available one might expect a 600 Mb/s interface to be perhaps three times as expensive as a 150 Mb/s interface and a 2.4 Gb/s interface to be perhaps ten times as expensive. For transmission systems spanning distances longer than 10 km, the costs increase, potentially doubling for distances in the 100–200 km range as one

[1] The costs used here are illustrative rather than precise, but have been selected to be typical of what one might expect to pay in the 1996–97 time frame as ATM technology becomes more widely available. The switching and transmission electronics numbers in particular, are somewhat on the low side, at the time of writing, but should be a fair reflection of typical costs in the near-term.

substitutes more powerful light sources and more sensitive receivers. At still longer distances, it becomes necessary to add amplifiers or repeaters periodically, with each such component having a cost roughly comparable to the transmission electronics that terminate the link.

- *Switching system cost.* The cost of a switch can be broken down into two parts, one which increases linearly with capacity and another which increases super-linearly. The linear term accounts for the per port ATM processing, while the super-linear term accounts for the switch's interconnection network. A typical cost function for a commercial ATM switch with n 150 Mb/s ports would be $\$200n + \$15n^2$ where the quadratic term is due to the use of a crossbar or bus. When $n = 16$ this gives a total cost of \$7,040 for the core switching function and \$15,040 when transmission interfaces are included. When $n = 64$, the cost becomes \$106,240. Newer switch architectures employing higher speed electronics and more efficient interconnection network designs can support n ports at 150 Mb/s (or higher speed ports with an equivalent total capacity) for a cost of $\$50n + \$10n \log_2 n$ [15]. A switch with this cost characteristic configured for 64 150 Mb/s ports would cost \$39,040, including transmission interfaces.

Consider now, three alternative designs for a campus network with 8,192 users with access links of 150 Mb/s, as shown in Figure 1. The hierarchical design, on the left, uses switches with 150 Mb/s links only, moderate size switches and a concentration ratio of 8-to-1 in the access switches (those at the bottom of the hierarchy). The middle design uses somewhat larger switches, uses 2.4 Gb/s links to the central hub and also uses a concentration ratio of 8-to-1. The third alternative is just a large central switch with 8,192 port interfaces. We can compare the costs of the alternative designs using the cost figures given above. We ignore the cost of the workstation and its network interface card, since this is the same in all three systems. For the hierarchical design, the fiber cost is \$9 per user, assuming that the access switches are distributed out near the users, the other switches are centralized and it take 250 meters of fiber to connect an access switch to this central location (we neglect the short lengths of fiber from the access switches to the end users). The transmission electronics for the hierarchical design costs \$750 per user and the switches cost \$126 per user (using the switch cost equation of $\$50n + \$10n \log_2 n$). For the "snow flake" network (the middle case), the fiber costs drop to about \$1 per user, the transmission interface costs to \$578 per user and the switch costs to \$118 per user. The star design has a substantially higher fiber cost of \$75 per user, but its transmission electronics cost is just \$500 per user and the switch costs are \$150 per user. The totals for the three designs are \$885 per user, \$697 and \$725 respectively. While the last two differ by only a few percent in cost, there is a clear performance advantage to the centralized switch, since it does not impose the 8-to-1 concentration ratio that is present in the other design. The situation for campus networks contrasts strongly with that for a geographically distributed network. For example, simply changing the 250 m distance to 25 km changes the fiber costs in the three cases to \$900, \$100 and \$7,500 per user, making the centralized switch the clear loser and the snow flake the clear winner.

This paper has six sections. Section 2 introduces the necessary definitions, formalizes the network design problem considered in this paper, and briefly summarizes our main

results. Section 3 addresses the computational complexity of the problem, and shows that the problem is NP-Complete. Section 4 describes our approximation techniques. Section 5 describes an optimal network for the special case of unit link costs. Section 6 provides some closing remarks and discussion of some of the issues we have neglected here.

2. Our Network Model

We describe a network by a complete digraph, $G = (V, E)$, where each vertex represents a *switch* and each directed edge represents a *link group*, comprising one or more physical transmission links. The vertices and edges of G have the following parameters associated with them:

- Each vertex u has an integer *source capacity* $\alpha(u)$, and an integer *sink capacity* $\omega(u)$, representing the maximum traffic rate that can originate or terminate at u.
- Each vertex pair (u, v) has a function $\gamma_\ell(u, v, x)$ representing the cost of constructing a link of capacity x from u to v.[2]

We also have a switch cost function $\gamma_s(x)$ giving the cost of a switch of total capacity x. If we assign a capacity $\kappa_\ell(u, v)$ to every edge (u, v) the resulting network cost is defined as

$$\sum_{(u,v)\in E} \gamma_\ell(u, v, \kappa_\ell(u, v)) + \sum_{u\in V} \gamma_s(\kappa_s(u)), \tag{1}$$

where $\kappa_s(u)$ is the capacity of switch u in the network and it equals the sum of the capacities of the links connecting it to other switches and the capacity of the connections to the end-systems it hosts (note that the capacity of the connections to the end-systems at a switch u may exceed $\alpha(u) + \omega(u)$, due to local communication). This model does not constrain traffic on a switch-pair basis, but just the total traffic at individual switches. This data is more readily available to a network designer and allows greater flexibility in the traffic that resulting networks are able to support. In this paper, we focus on this version of the problem, but in Section 6 we show how our model can be generalized to allow the specification of quite general traffic constraints, while preserving the worst-case viewpoint that we advocate.

In order to define the notion of nonblocking networks, we first need to define connection requests and their routing in the network. A *connection request* $R = (S, D, w)$ comprises a non-empty set of *sources* S, a non-empty set of *destinations* D and an integer *weight* $w \leq B$, where B is a maximum connection weight. If $S = D$ we say the request is *symmetric* and if $|S \cup D| = 2$, we say the request is *point-to-point*. A *route* T for a request R is a subgraph of G for which the underlying undirected graph is a tree and in which there is a directed path from every vertex in S to every vertex in D. A collection of routes C places a connection weight $\lambda_C(u, v)$ on an edge (u, v), which is defined as the sum of the weights of all routes

[2] We require that the costs satisfy the *triangle inequality*, meaning that the direct path of any given capacity between two vertices is never more expensive than an indirect path with the same capacity.

that include the edge (u, v). $\lambda_C(u)$ denotes the weight on a switch u, which is equal to the sum of the weights of its incident edges.

A set of connection requests is *valid* if, for every vertex u, the sum of the weights of the requests containing u in their source and sink sets, respectively, does not exceed $\alpha(u)$ and $\omega(u)$. A collection of routes C is *valid* if it satisfies a set of valid connection requests, and if $\lambda_C(u, v) \leq \kappa_\ell(u, v)$, for every edge (u, v).

A *state* of a network is a valid set of routes. A *routing algorithm* is a procedure that maintains a valid set of routes under the following four operations: (1) add a new route satisfying a specified connection request; (2) remove an existing route; (3) add a new vertex to either the source set, the destination set, or both for some route in the current state; (4) remove some vertex from either the source set, the destination set, or both for some route in the current state.[3] We are only concerned with routing algorithms that are *incremental*, meaning that they only add, delete or modify a single route when carrying out a requested operation and that they cannot both add and remove edges from an existing route in a single operation.

The *reachable states* for a routing algorithm on a network with specified link capacities is the set of all states that can be reached by sequences of the four operations given above, starting from the empty state. We say that a network is *nonblocking* under a given routing algorithm if for every reachable state and every operation request whose completion would not exceed the source or sink capacity of any vertex in that state, the algorithm produces a new state satisfying the operation request.

The use of a scalar to represent the resource requirements of a connection is clearly a simplification. In the ATM context, our model applies directly to *constant bit rate* applications and applications whose resource requirements can be adequately described by an *effective data rate*. It does not apply as well to applications which have no specific rate requirement, but adjust their transmission rate dynamically to take advantage to available resources or in reaction to congestion. However, we argue that networks designed to be nonblocking for applications whose resource requirements can be adequately described by a scalar value will provide good performance for those more dynamic applications as well.

The *nonblocking network design problem* is to *determine a set of link capacities* that will yield a nonblocking network of least cost under either a specified routing algorithm or some routing algorithm from a specified class of routing algorithms. In the latter case, the design problem is to produce both the link capacities and a specific routing algorithm from the given class, for which the network is nonblocking. Figure 2 shows an instance of the network design problem on the left and a solution on the right. On the left, the numbers next to each vertex denote the switch capacities, $\alpha(v), \omega(v)$; the number next to each edge denotes the link cost per unit capacity (assuming symmetric link costs). The solution on the right shows directional capacities on links. This network is nonblocking if connections are always routed using shortest available paths.

In many situations, some special cases of the network design problem are of interest. In the *linear* cost version, switch costs are zero and all link costs satisfy $\gamma_\ell(u, v, x) =$

[3] A routing algorithm may fail to carry out operations of type (1) or (3), but will always carry out operations of type (2) or (4).

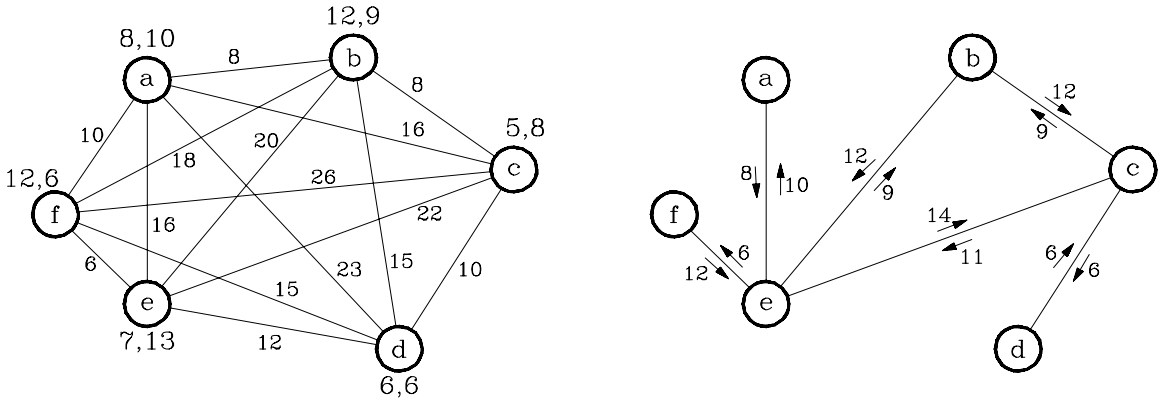

Figure 2: An example of the design problem and a suboptimal solution

$x \times \gamma(u,v)$, where $\gamma(u,v)$ is a constant that depends only on u and v. The *symmetric* version of the problem has $\alpha(u) = \omega(u)$ for all vertices u, $\gamma_\ell(u,v,x) = \gamma_\ell(v,u,x)$ for all pairs u, v and restricts the choice of link capacities so that $\kappa_\ell(u,v) = \kappa_\ell(v,u)$. In the *balanced* version of the problem, we have $\sum_{u \in V} \alpha(u) = \sum_{u \in V} \omega(u)$.

2.1. Summary of Results

In this paper, we focus on the linear link cost model, that is, $\gamma_\ell(u,v,x) = x \times \gamma(u,v)$ and switch costs are zero. While this is clearly an idealization of reality, it is surprisingly accurate in many common situations. As we saw in the campus network example earlier, link costs are not strictly linear, since the cost of the transmission electronics for a 2.4 Gb/s link is less than 16 times the cost of the transmission electronics for a 150 Mb/s link, while the fiber costs for the 2.4 Gb/s link are essentially the same as that at 150 Mb/s. However, once we move to link groups with capacities above 2.4 Gb/s, we can only obtain additional capacity (today) by operating multiple links in parallel, implying that this portion of the capacity-cost curve has an essentially linear growth characteristic. In large scale networks, we believe link groups with multiple parallel links will be the norm, rather than the exception and where this is true the linear cost assumption is a reasonable one to make.

The assumption of zero switch costs is also reasonable in the context of switches with efficient architectures, such as [15]. For example, a switch with a cost characteristic of $\$20n + \$10n \log_2 n$ (where n is the number of 150 Mb/s interfaces it can support) has a cost of $\$110n$ when $n = 64$ and $\$180n$ when $n = 65,536$. If we don't represent switch costs explicitly in our model, but add \$145 to the cost of the transmission electronics as a way of approximately accounting for the switch costs (this changes the allocated cost of the transmission electronics for a 150 Mb/s link from \$500 to \$645), the error in the calculated cost of a network with short links is less than 6%, so long as we constrain ourselves to switches for which n is in the range 64 to 65,536. For networks with longer links the relative error diminishes. With this understanding, the linear link cost model

is very useful in obtaining insight into the topological characteristics of optimal and near optimal networks.

For the linear link costs model, we prove that networks with a star topology achieve near-optimal cost. In particular, for the symmetric case, we prove that the least cost nonblocking network of *arbitrary* topology has cost at least half the cost of the cheapest nonblocking star network. The ratio becomes 1/3 when the source and sink traffic capacities are asymmetric, but balanced. For arbitrary traffic capacities, the performance ratio of the star networks degrades gracefully (cf. Theorem 4.1). Finally, we show that in the special case of *unit* link cost function, meaning $\gamma(u, v, x) = c \cdot x$ for some absolute constant c, a star network is indeed optimal. This last case, while apparently a gross simplification of the problem applies quite well to campus networks, since, as the example given earlier shows, when the distances are short enough, the costs of links are determined almost entirely by the electronics, not the fiber. The surprising efficiency of star networks in these situations has an important consequence: a star network (or indeed any tree-structured network) that is nonblocking for point-to-point channels is also nonblocking for multicast channels (assuming that all the switches are).

Even in the linear link cost model, the problem of computing a least-cost nonblocking network turns out to be NP-Complete, meaning that approximation algorithms are the only recourse for designing provably cost-effective nonblocking networks. We give several hardness results in the next section, setting the stage for our analysis of the cost-effectiveness of the star networks.

3. Computational Complexity of the Problem

A solution to the network design problem asks for a cheapest set of link capacities as well as an incremental strategy for setting up valid connections. In general, the routing problem in itself is a hard problem. There is a variety of routing strategies one can employ. One of the simplest is *fixed path routing* in which we always use a particular path to route between a given pair of vertices (this can also be generalized to multicast channels). Fixed path routing is the only real choice for tree-structured networks, and while not generally used in networks with more complex topology, it offers a useful point of comparison, even in these cases. Most commonly used routing algorithms are some variant on *shortest available path routing* in which a request to connect a given pair of destinations uses the shortest path with sufficient unused capacity (again, this can be generalized to multicast channels). It's usual to impose an efficiency constraint that rules out the use of paths that use an "unreasonable" amount of resources. The following theorem, proved in [4], shows that determining whether a given network is nonblocking is NP-Hard.

THEOREM 3.1. [4] *Let V be a set of switches, let $\alpha(u), \omega(u)$ be their source and sink capacities, and let $\kappa_\ell(u, v)$ be the capacity of link (u, v). Suppose that all connection requests have weights that are multiples of a minimum weight b and connections are routed using the shortest available path. Then the problem of deciding whether the network is nonblocking for point-to-point connection requests is NP-Hard in the strong sense.*

The intractability of *checking* whether a network is nonblocking does not imply that *designing* one is also hard. However, we can show that several simple versions of the design problem are indeed intractable. In the first theorem, we let the source and sink capacities be arbitrary, with no constraints of symmetry or balance. In this version, the well-known Steiner tree problem in graphs with edge costs in $\{1, 2\}$ turns out to be special case of the network design problem. (Let $G = (V, E)$ be a graph, $w(e) \in \{1, 2\}$ be weights on edges, $R \subset V$ be a subset, and B a positive integer bound. The Steiner Tree problems asks if there is a subtree of G spanning all nodes of R with a total cost of at most B. This version of the Steiner Tree problem was proved MAXSNP-Hard by Bern and Plassman [2]; see also [6].)

THEOREM 3.2. *Given a set of switches V, their source and sink capacities $\alpha(v)$, $\omega(v)$, and a linear link cost function $\gamma(u, v)$ for each switch-pair in V, the problem of finding a minimum cost, nonblocking network for $(V, \alpha, \omega, \gamma)$ is MAXSNP-Hard.*

PROOF. The set of switches V is the set of nodes V. The link costs are the same as the edge costs in G, namely, $\gamma(u, v) = w(u, v)$. Observe that the link costs satisfy triangle inequality. We pick an arbitrary "root" node $r \in R$, from the Steiner subset $R \subset V$. Set $\alpha(r) = 1$ and $\omega(r) = 0$. For the remaining Steiner nodes $u \in R$, we set $\alpha(u) = 0$ and $\omega(u) = 1$. All other nodes of V have $\alpha(v) = \omega(v) = 0$, where $v \in V - R$. Thus, the root node can originate one unit of traffic, but it has no termination capacity. Every other node of the Steiner subset R has one unit of termination capacity and no origination capacity. The nodes in $V - R$ have no origination/termination capacity at all.

Let $\mathcal{N}^\star$ be a minimum cost nonblocking network for the above instance. It is easily seen that V admits a Steiner Tree of cost B on R if and only if $cost(\mathcal{N}^\star) \leq B$. In particular, every Steiner tree spanning R can be turned into a nonblocking network, by directing all edges away from the root node and assigning unit capacity to each link. Conversely, every nonblocking network with cost $\leq B$ can be converted to a Steiner tree of with cost $\leq B$. □

We can also show that the network design problem even with symmetric capacities is hard, for a slight variation of the link cost model. In particular, assume that setting up a link from u to v of capacity $\kappa_\ell(u, v)$ has cost

$$c(u, v) \; + \; \gamma(u, v) \times \kappa_\ell(u, v),$$

where $c(u, v)$ is a fixed installation cost, independent of the link capacity. We can show that this version of the problem is NP-complete by way of a polynomial time reduction from the well-known *set cover problem*:

> Given a finite set X and a family $\mathcal{F} = \{S_1, S_2, \ldots, S_m\}$ of subsets of X, find a *minimum cardinality* subset $J \subseteq \{1, 2, \ldots, m\}$ such that $\cup_{j \in J} S_j \; = \; X$.

The proof is omitted due to space limitations. In view of these hardness results, we focus our attention, in the following section, on efficient algorithms for designing nonblocking networks of provably small cost.

4. Designing Low-Cost Nonblocking Networks

We show that *star networks* produce nearly optimal results. In particular, we prove that there exists a star network, rooted at one of the nodes of V, that is nonblocking and has a cost at most twice the minimum cost in the symmetric case (i.e., $\alpha(v) = \omega(v)$ for all v). In the balanced case, the same network is also shown to be within a factor 3 of optimal. As the balance condition worsens, the quality of approximation degrades gracefully: we prove that there is a star network with cost no more than $2 + \frac{\sum \alpha(u)}{\sum \omega(u)}$ times the optimal, where we assume without loss of generality that $\sum \alpha(u) \geq \sum \omega(u)$. An optimal nonblocking star can be found algorithmically in $O(n^2)$ time, where n is the number of switches.

We will bound the cost of an optimal star network in terms of a quantity $\mathcal{D}$ defined below, and then derive a lower bound on the cost of a cheapest nonblocking network also in terms of $\mathcal{D}$ to establish our results. We will frequently need to refer to the total source and sink capacities. For convenience, let us introduce the following shorthand notation:

$$\mathcal{A} = \sum_{v \in V} \alpha(v) \qquad \text{and} \qquad \mathcal{Z} = \sum_{v \in V} \omega(v).$$

Throughout the following discussion, we assume without loss of generality that $\mathcal{A} \geq \mathcal{Z}$. The quantity $\mathcal{D}$ is defined as follows:

$$\mathcal{D} = \sum_{u} \sum_{v} \alpha(u) \times \omega(v) \times \gamma(u, v). \tag{2}$$

We are now ready to proceed with our proof of the approximation bound; we first establish the general upper bound, and then sharpen it further for the symmetric case of switch capacities.

4.1. General Switch Capacities

In establishing the upper bound, we use an intermediate network that has the form of a double star. The double star $S(v_k, v_l)$ corresponding to an ordered pair (v_k, v_l) is defined by the following link capacities:

1. $\kappa(v_i, v_l) = \alpha(v_i)$, for $i \neq l$;
2. $\kappa(v_k, v_i) = \omega(v_i)$, for $i \neq k$;
3. $\kappa(v_l, v_k) = \mathcal{Z}$.

All other links in $S(v_k, v_l)$ have zero capacity. See Figure 3 for an illustration. We will show that the cheapest double star achieves the desired cost, but first let us show that the double star described above is indeed a nonblocking network.

LEMMA 4.1. *The double star $S(v_k, v_l)$ is a nonblocking network for $(V, \alpha, \omega, \gamma)$.*

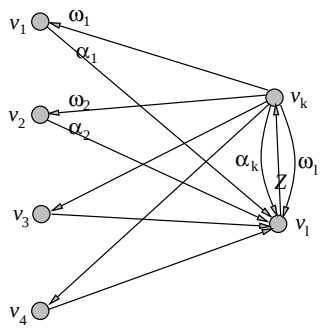

Figure 3: Illustrating a double star.

PROOF. The link (v_l, v_k) clearly has sufficient bandwidth to route all valid connections, since the maximum traffic to all receiving switches, other than v_l itself, cannot exceed $\mathcal{Z} - \omega(v_l)$. Since each v_i has outgoing link capacity $\alpha(v_i)$ and each v_j has incoming link capacity $\omega(v_j)$, it is easily seen that no valid connection request is blocked. □

In order to complete our proof of the approximation bound, we show below that there exists a double-star in $B(V)$ whose cost is within a factor $2 + \frac{\mathcal{A}}{\mathcal{Z}}$ of the cost of an optimal network.

LEMMA 4.2. *A minimum cost double star of V has cost no greater than*

$$\frac{(\mathcal{A} + 2\mathcal{Z})\mathcal{D}}{\mathcal{A} \cdot \mathcal{Z}}.$$

PROOF. We prove the lemma by considering a multiset of double stars of V and arguing that the cost of an *average* double star in this multiset has the claimed bound. Since the minimum of a set cannot exceed its average, the lemma follows. So, let $\mathcal{M}$ denote the multiset of double stars, in which $S(v_k, v_l)$ appears $\alpha(v_k) \times \omega(v_l)$ times. The family $\mathcal{M}$ has size

$$\begin{aligned} |\mathcal{M}| &= \sum_{k=1}^{|V|} \sum_{l=1}^{|V|} \alpha(v_k) \times \omega(v_l) \\ &= \mathcal{A} \times \mathcal{Z}. \end{aligned} \tag{3}$$

Let us now count the total cost of all the double stars in this multiset. We do this by counting the contribution of each edge (v_k, v_l), and summing over all pairs. An edge (v_k, v_l) contributes costs in three ways:

1. In the double star $S(v_l, v_k)$, (v_k, v_l) has capacity $\mathcal{Z}$. This double star appears $\alpha(v_l) \times \omega(v_k)$ times, and, by symmetry of the link cost, $\gamma(v_k, v_l) = \gamma(v_l, v_k)$. Thus, the the total contribution is
$$\alpha(v_l)\omega(v_k)\gamma(v_k, v_l)\mathcal{Z}.$$

2. In each of the double stars $S(v_k, v_j)$, it appears with capacity $\omega(v_l)$. The total number of these double stars in $\mathcal{M}$ is $\alpha(v_k) \times \sum_{j=1}^{|V|} \omega(v_j)$, which implies that the total contribution is at most
$$\alpha(v_k)\omega(v_l)\gamma(v_k, v_l) \times \mathcal{Z}$$

3. In each of the double stars $S(v_i, v_l)$, it appears with capacity $\alpha(v_k)$. The total number of these double stars in $\mathcal{M}$ is $\omega(v_l) \times \sum_{i=1}^{|V|} \alpha(v_i)$, which implies that the total contribution is at most
$$\alpha(v_k)\omega(v_l)\gamma(v_k, v_l) \times \mathcal{A}.$$

Recalling that $\mathcal{D} = \sum_k \sum_l \alpha(v_k)\omega(v_l)\gamma(v_k, v_l)$, we obtain that the total cost of all the double stars in $\mathcal{M}$ is at most
$$(\mathcal{A} + 2\mathcal{Z})\mathcal{D}. \tag{4}$$

Thus, we get an upper bound on the cost of an average double star in $\mathcal{M}$ by dividing the quantity in Eq. (4) by the quantity in Eq. (3), which gives the bound claimed in the lemma. This completes the proof. □

Finally, we show that triangle inequality implies that the cost of a cheapest nonblocking star cannot exceed the cost of a cheapest double star. In particular, we show that the double star $S(v_k, v_l)$ can be converted to a star rooted at v_l with no increase in cost. In the double star $S(v_k, v_l)$, we leave all incoming links of v_l the same, but transfer all outgoing links of v_k to v_l. Clearly, this yields a nonblocking star rooted at v_l. The following lemma proves the bound on the cost.

LEMMA 4.3. *The cost of the cheapest nonblocking star rooted at v_l does not exceed the cost of $S(v_k, v_l)$.*

PROOF. In modifying the double star into the star network, we effectively replace the path (v_l, v_k, v_i) with the direct path (v_l, v_i). By triangle inequality,

$$(\mathcal{Z} - \omega(v_k)) \times \gamma(v_l, v_k) \; + \; \sum_{i \neq k} \omega(v_i) \times \gamma(v_k, v_i) \;\; \geq \;\; \sum_{i \neq k} \omega(v_i) \times \gamma(v_l, v_i).$$

Consequently, the cost of the double star is at least equal to the cost of the star rooted at l. This completes the proof. □

4.2. An Improved Bound for Symmetric Switch Capacities

In this case, we can directly bound the cost of an optimal star network. Consider the least cost nonblocking star rooted at node u and note that it has cost

$$\sum_{v \neq u} (\alpha(v) + \omega(v))\gamma(u, v).$$

Let $\mathcal{M}$ denote the multiset of stars in which, for every u, the star with root u appears exactly $\alpha(u)$ times. The cost of all the stars in $\mathcal{M}$ is thus

$$\sum_u \alpha(u) \sum_{v \neq u} (\alpha(v) + \omega(v))\gamma(u,v) = \sum_u \sum_v \alpha(u)(2\omega(v))\gamma(u,v) = 2\mathcal{D},$$

where we've used the fact that $\alpha(v) = \omega(v)$. Since $|\mathcal{M}| = \mathcal{A}$, it follows that the cheapest star in $\mathcal{M}$ has cost no more than $2\mathcal{D}/\mathcal{A}$, giving the following lemma.

LEMMA 4.4. *Let V be a set of switches, with symmetric source and sink capacities, $\alpha(v) = \omega(v)$, and link costs $\gamma(u,v)$ for all switch- pairs u, v. Then, the cost of a cheapest nonblocking star network for $(V, \alpha, \omega, \gamma)$ is at most $\frac{2\mathcal{D}}{\mathcal{A}}$.*

In order to show that these star networks are near optimal, we need to establish a lower bound on the cost of any nonblocking network. We do this in the following subsection.

4.3. A Lower Bound on the Cost of an Optimal Network

Suppose $\mathcal{N}^\star$ is a nonblocking network for the switch capacities $\alpha(v)$, $\omega(v)$ and link costs $\gamma(u,v)$, where $u, v \in V$. Being a nonblocking network, $\mathcal{N}^\star$ is able to route any set of switch capacity-compliant connections. Consider a feasible connection between u and v at data rate $f(u,v)$, where feasibility dictates that $f(u,v) \leq \min\{\alpha(u), \omega(v)\}$. Then, by triangle inequality, the route(s) used by $\mathcal{N}^\star$ to set up this connection must cost at least $\gamma(u,v) \times f(u,v)$. Now, if there are two simultaneously feasible connections, one from u to v at rate $f(u,v)$ and another from x and y at rate $f(x,y)$, then the *linearity* of link costs implies that the network has cost at least

$$\gamma(u,v) \times f(u,v) + \gamma(u,v) \times f(u,v). \tag{5}$$

Thus, any set of simultaneously feasible connections implies a lower bound of the form Eq. (5) on the cost of $\mathcal{N}^\star$. In order to get the best lower bound, we seek connections of maximum cost.

The problem of finding a set of simultaneous connections maximizing the cost can be re-cast as a *maximum-weighted matching problem*. To do this, we first carry out a *node-splitting transformation*, which splits a node u into $\alpha(u)$ source nodes and $\omega(u)$ sink nodes, each with unit capacity. More formally, let $v_i \in V$ be a switch with source capacity $\alpha_i = \alpha(v_i)$ and sink capacity $\omega_i = \omega(v_i)$. We replace v_i with α_i copies of itself labeled *source nodes* $a_{i1}, a_{i2}, \ldots, a_{i\alpha_i}$, and with ω_i copies labeled *sink nodes* $z_{i1}, z_{i2}, \ldots, z_{i\omega_i}$. Assign $\alpha(a_{ij}) = 1$ and $\omega(a_{ij}) = 0$, and $\alpha(z_{ij}) = 0$ and $\omega(z_{ij}) = 1$. Thus, each source node has send capacity of one and receive capacity of zero, while each sink node has the send capacity of zero and receive capacity of one. Now, construct a bipartite graph by joining each a-node to each z-node and "inheriting" the link cost from the original problem. Specifically, we assign

$$\gamma(a_{ij}, z_{kl}) = \gamma(v_i, v_k), \qquad \text{for } j = 1, 2, \ldots, \alpha_i, \text{ and } l = 1, 2, \ldots, \omega_i.$$

An example of our graph transformation is shown in Figure 4. We call this bipartite graph $B(V)$. Observe that $B(V)$ has $\mathcal{A} + \mathcal{Z}$ nodes and $\mathcal{A} \times \mathcal{Z}$ edges, where recall that $\mathcal{A} = \sum_v \alpha(v)$ and $\mathcal{Z} = \sum_v \omega(v)$, and we assume that $\mathcal{A} \geq \mathcal{Z}$. In order to simplify the notation, let us renumber the nodes so that the source nodes are labeled $a_1, a_2, \ldots, a_{\mathcal{A}}$, and the sink nodes are labeled $z_1, z_2, \ldots, z_{\mathcal{Z}}$.

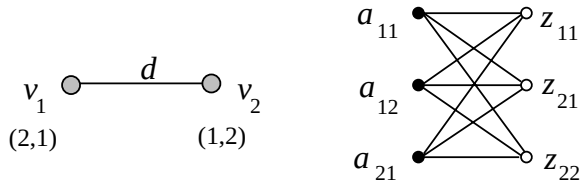

Figure 4: Illustrating the graph transformation. In the figure, $\alpha(v_1) = 2$, $\omega(v_1) = 1$, and $\alpha(v_2) = 1$, $\omega(v_2) = 2$. Edges (a_{11}, z_{11}), (a_{12}, z_{11}), (a_{21}, z_{11}), and (a_{21}, z_{22}) have link costs zero; others have cost d.

Let M denote an arbitrary matching in $B(V)$ (recall that a matching is a collection of vertex disjoint edges). We claim that a maximum-weight matching in $B(V)$ has weight at least $\mathcal{D}/\mathcal{A}$, where the weight of the matching is the total cost of its edges.

LEMMA 4.5. *Let M be a maximum-weight matching in $B(V)$. Then, cost$(M) \geq \mathcal{D}/\mathcal{A}$.*

PROOF. First, observe that

$$\mathcal{D} = \sum_{i=1}^{\mathcal{A}} \sum_{j=1}^{\mathcal{Z}} \gamma(a_i, z_j);$$

this follows because the node-splitting transformation makes $\alpha(u) \times \omega(v)$ copies of the edge (u, v). The number of different matching in $B(V)$ is $\binom{\mathcal{A}}{\mathcal{Z}} \cdot \mathcal{Z}!$. (The first term counts the number of ways to pick which $\mathcal{Z}$ source nodes to match with the sink nodes, and the second term counts the number of ways to do this matching.) Every edge of $B(V)$ gets counted $\binom{\mathcal{A}-1}{\mathcal{Z}-1} \cdot (\mathcal{Z}-1)!$ times over all the matchings. Thus, the total weight of all the matchings is $\binom{\mathcal{A}-1}{\mathcal{Z}-1} \cdot (\mathcal{Z}-1)! \cdot \mathcal{D}$. Since the maximum of a set is at least as large as its average, the maximum-weight matching satisfies

$$\begin{aligned} cost(M) &\geq \frac{\binom{\mathcal{A}-1}{\mathcal{Z}-1} \cdot (\mathcal{Z}-1)! \cdot \mathcal{D}}{\binom{\mathcal{A}}{\mathcal{Z}} \cdot \mathcal{Z}!} \\ &= \frac{\mathcal{D}}{\mathcal{A}}, \end{aligned}$$

which completes the proof. □

A matching in $B(V)$ corresponds (uniquely) to a set of valid connections, with the same total cost as the matching, giving the following corollary.

COROLLARY 4.1. *Let V be a set of switches with source and sink capacities $\alpha(v), \omega(v)$, for $v \in V$, and assume that the link cost $\gamma(u,v)$, for all switch-pairs $(u,v) \in V \times V$, satisfies the triangle inequality. Then, a minimum-cost nonblocking network for $(V, \alpha, \omega, \gamma)$ has cost at least $\frac{\mathcal{D}}{\mathcal{A}}$.*

4.4. Approximation Ratios for Star Networks

Comparing the cost of a cheapest star network to the lower bound of Corollary 4.1, we can bound the approximation factor of our star network. The approximation factor is given by

$$\begin{aligned} \frac{cost(\text{cheapest star})}{cost(\mathcal{N}^\star)} &\leq \frac{\frac{(\mathcal{A}+2\mathcal{Z})\mathcal{D}}{\mathcal{A}\cdot\mathcal{Z}}}{\frac{\mathcal{D}}{\mathcal{A}}} \\ &\leq \frac{\mathcal{A}+2\mathcal{Z}}{\mathcal{Z}} \\ &\leq 2+\frac{\mathcal{A}}{\mathcal{Z}} \\ &\leq 3 \quad \text{if } \mathcal{A}=\mathcal{Z}. \end{aligned}$$

Thus, in the balanced case, namely $\mathcal{A} = \mathcal{Z}$, there exists a nonblocking star network for the network design problem $(V, \alpha, \omega, \gamma)$ whose cost does not exceed three times the cost of an optimal network. Without any balance condition, the cost of the best star network is within $2 + \frac{\mathcal{A}}{\mathcal{Z}}$ times of the optimal. For the symmetric capacity case, the ratio of the star to optimal network is 2 (cf. Lemma 4.4). We conclude with the following theorem.

THEOREM 4.1. *Let V be a set of switches, with source and sink capacities $\alpha(v)$ and $\omega(v)$, and link costs $\gamma(u,v)$ for all switch pairs u, v. Then, the ratio between the cost of a cheapest nonblocking star and an optimal network is at most 2 if the switch capacities are symmetric, at most 3 if the switch capacities are balanced, and at most $2 + \frac{\mathcal{A}}{\mathcal{Z}}$ in general, where $\mathcal{A} \geq \mathcal{Z}$.*

5. Unit Link Costs

We now consider the case when all link costs are the same, and show that a star network is optimal when the switch capacities are balanced. Despite being an idealized case, it applies to practical situations where the link costs are dominated by the cost of the terminating electronics, which is generally the case within campus networks. Since all links have the same cost, without loss of generality, we assume that $\gamma(u,v) = 1$, for all u, v. In this case, the problem can be specified with three parameters: (V, α, ω). We first prove the following lemma, which is useful in the proof of the main theorem.

LEMMA 5.1. *Suppose V is a set of switches, with balanced source and sink capacities $\alpha(v)$ and $\omega(v)$, and unit link cost function between pairs of switches. Let $\mathcal{N}$ be a nonblocking*

network for (V, α, ω) *such that* $\kappa_\ell(u,v) \geq \min\{\alpha(u), \omega(v)\}$, *for all* $(u,v) \in V \times V$. *Then, the following holds:*

$$cost(\mathcal{N}) \geq \sum_v (\alpha(v) + \omega(v)) - \max_v (\alpha(v) + \omega(v)).$$

PROOF. We note that the bound on the right hand side is the cost of a nonblocking star, rooted at the node with a maximum capacity; unit link costs imply that the cost of a network equals its total link capacity. In counting the link capacities in $\mathcal{N}$, we charge each link to its *destination* node. Let v_m denote the switch with the maximum capacity (source or sink) of all switches, and without loss of generality assume that

$$\omega(v_m) = \max_{v \in V}\{\alpha(v), \omega(v)\}.$$

Considering any other node v_i, where $i \neq m$, we get

$$\kappa_\ell(v_i, v_m) \geq \alpha(v_i), \tag{6}$$

since $\alpha(v_i) \leq \omega(v_m)$. All these links are charged to v_m, and they sum to $\sum_v \alpha(v) - \alpha(v_m)$.

Next, if the total incoming link capacity at each v_i, for $i \neq m$, is at least $\omega(v_i)$, then we get the desired bound on the overall cost of the network, completing the proof. So, assume that the incoming link capacity falls short at some node, say, v_i. Since we must have $\kappa_\ell(v_j, v_i) \geq \min\{\alpha(v_j), \omega(v_i)\}$, for all $v_j \neq v_i$, the total incoming link capacity at v_i fails to add up to $\omega(v_i)$ only if the following holds:

$$\omega(v_i) > \sum_v \alpha(v) - \alpha(v_i);$$

that is, the sink capacity of v_i exceeds the combined source capacity of all other nodes. When this happens, we conclude that

$$\alpha(v_i) + \omega(v_i) > \sum_{v \in V} \alpha(v). \tag{7}$$

We now re-apply the argument, using v_i in place of v_m as the purported root of the star. Since $\omega(v_i) > \sum_{v \in V} \alpha(v) - \alpha(v_i)$, it follows that each *incoming* link (v_k, v_i) at v_i has capacity $\kappa_\ell(v_k, v_i) = \alpha(v_k)$. We charge these links to v_i, and consider the incoming links at any other node v_k. Can it happen again that at some node v_k, for $k \neq i$, we find

$$\omega(v_k) > \sum_v \alpha(v) - \alpha(v_k)? \tag{8}$$

Suppose it did. Then, inequalities (7) and (8) together imply that

$$\begin{aligned} (\alpha(v_i) + \omega(v_i)) + (\alpha(v_k) + \omega(v_k)) &> 2 \sum_{v \in V} \alpha(v) \\ &= \sum_{v \in V} (\alpha(v) + \omega(v)), \end{aligned} \tag{9}$$

which is clearly not possible. Thus, the incoming links at each node v_k, for $i \neq k$, sum to $\omega(v_k)$, and thus the total link capacity of $\mathcal{N}$ is at least

$$\sum_v (\alpha(v) + \omega(v)) \; - \; \max_v (\alpha(v) + \omega(v)),$$

and the proof is completed. □

THEOREM 5.1. *Let V be a set of switches, with source and sink capacities $\alpha(v)$ and $\omega(v)$, and assume unit link cost function between pairs of switches. Then, for balanced switch capacities, a minimum cost nonblocking star network is an optimal network.*

We can now prove the result that, for unit link costs, a star network is optimal.

THEOREM 5.2. *Let V be a set of switches, with source and sink capacities $\alpha(v)$ and $\omega(v)$, and assume unit link cost function between pairs of switches. Then, for balanced switch capacities, a minimum cost nonblocking star network is an optimal network.*

PROOF. We show that any nonblocking network must have a total link capacity at least

$$\sum_{v \in V} (\alpha(v) + \omega(v)) \; - \; \max_{v \in V} (\alpha(v) + \omega(v)) . \tag{10}$$

It is easy to see that this matches the cost of a cheapest nonblocking star network, obtained by choosing as root the switch with the maximum source plus sink capacity. Let $\mathcal{N}^\star$ be an optimal nonblocking network, and let $\kappa_\ell(u, v)$ denote the capacity of the link (u, v); if there is no link between u and v, this capacity is zero. Consider any pair of nodes $(u, v) \in V \times V$ for which the following inequality holds:

$$\kappa_\ell(u, v) < \min\{\alpha(u), \omega(v)\}. \tag{11}$$

We set up two connections from u to v, first at the rate of $\kappa_\ell(u, v)$, and second at the rate $f(u, v) = \min\{\alpha(u), \omega(v)\} - \kappa_\ell(u, v)$. Due to the capacity constraint, the second connection must use an indirect path, requiring at least two links. We now tear-down the *first* connection, freeing up switch capacities $\kappa_\ell(u, v)$ at both u and v.

Since connection rerouting is not permitted in nonblocking networks, the second connection continues to be routed along the indirect path. This connection consumes $f(u, v)$ units of source (resp. sink) capacity of u (resp. v). It also consumes at least $2f(u, v)$ link capacities in $\mathcal{N}^\star$, by virtue of being an indirect path. Subtract $f(u, v)$ from the switch capacities of u and v, and link capacities of all the links in the indirect path used by the connection. Observe that this modification keeps the switch capacities balanced. (In order not to introduce extra notation, we continue to use $\alpha(v)$, $\omega(v)$, and $\kappa_\ell(u, v)$ for the *residual* capacities of switches and links.)

We now repeat the connection setup procedure at any other link for which the condition in Ineq. (11) holds, until no such link exists. Suppose that the total source capacity consumed by the indirect connections is $\mathcal{A}_1$; an equal amount of sink capacity is also consumed.

By the simultaneous connection argument used in Eq. (5), these (indirect) connections saturate at least

$$2\mathcal{A}_1 \tag{12}$$

units of link capacity in $\mathcal{N}^\star$.

When the condition in (11) no longer holds, every node-pair $(u, v) \in V \times V$ satisfies:

$$\kappa_\ell(u, v) \;\geq\; \min\{\alpha(u), \omega(v)\},$$

and the total residual source capacity is $\mathcal{A} - \mathcal{A}_1$. We now invoke Lemma 5.1 on the residual network, which must be nonblocking for the residual (balanced) capacities. Combining the lower bound of Lemma 5.1 with Eq. (12), we conclude that

$$\begin{aligned} cost(\mathcal{N}^\star) &\geq 2\mathcal{A}_1 + 2(\mathcal{A} - \mathcal{A}_1) - \max_v \left(\alpha(v) + \omega(v)\right) \\ &\geq \sum_v \left(\alpha(v) + \omega(v)\right) - \max_v \left(\alpha(v) + \omega(v)\right), \end{aligned}$$

which completes the proof. □

6. Discussion and Future Research Directions

The results reported here represent first steps to developing an understanding of how best to design ATM networks. We feel strongly that the worst-case viewpoint we have adopted is very productive and has already yielded important insights. At the same time, our abstract formulation of the network design problem ignores certain aspects of the real problem that need to be addressed. The biggest single limitation of the formulation of the problem given here is its inability to express more 'fine-grained' constraints on the traffic than can be captured through source and sink capacities of individual switches. While this may not be a serious problem in campus networks in which the natural "communities of interest" share common switches, it does not allow us to model more general situations, meaning that we miss opportunities to obtain more economical designs. In [4], the per-switch source/sink capacities are extended to allow source and sink capacities for clusters of switches and "clusters-of-clusters," leading to a hierarchy of traffic constraints. This allows network planners to specify traffic patterns that follow natural hierarchical patterns (often found in large organizations, for example). For such traffic, we refine our definition of nonblocking networks to yield networks that are nonblocking so long as the traffic stays within the specified constraints. [4] gives simulation results showing that tree structured networks that follow the traffic clustering in the natural way and are configured to be nonblocking are close to optimal.

However, hierarchical clustering is only one of the natural patterns exhibited by network traffic. In networks spanning large geographic distances, there is a natural tendency for reduced communication beyond a certain point. We can incorporate distance-related traffic constraints within our framework in a straightforward way. The simplest way to do this is to define pairwise distances among all the switches and to associate a second source and

sink capacity with each switch that specifies the maximum traffic originating or terminating at the switch that goes to switches within some critical distance d. We have found that when distance constraints are included, the best networks are no longer tree-structured, but have more complex topologies. One class of designs that we have considered (based on a uniform triangulation of the plane) produces networks in which the cost for the portion of the network that carries "local traffic" is at most

$$2d \sum_u (\alpha_d(u) + \omega_d(u))$$

where $\alpha_d(u)$ and $\omega_d(u)$ represent the local source/sink capacities and where switches are assigned a location in the Euclidean plane with inter-switch distances given by the Euclidean distances. While this can be far from optimal, in the worst-case (for example, consider a problem instance in which no switch pairs are within distance d of one another), we conjecture that for typical situations, any nonblocking network for this case will have a cost of at least $d \sum_u (\alpha_d(u) + \omega_d(u))$ for the local traffic.

To allow network planners to easily express natural traffic patterns without requiring the specification of an excessive amount of information, we introduce general traffic constraints of the form $\mu(S_1, S_2)$ which specifies an upper bound on the total traffic from a set of switches S_1 to another set S_2. Using constraints of this form, a network planner can express hierarchical clustering or simple distance constraints as well as less regular constraints on traffic caused by natural geographic or cultural barriers (such as a mountain range, or a language difference). Planners need only specify constraints for those pairs of sets where it's useful to do so. In general, the more information given, the more closely a network design can be tailored to real needs, yielding lower overall cost.

An important feature of our approach is that given any instance of the network design problem, including arbitrary node-set pair constraints, there is a straightforward method for computing a lower bound on the cost of any nonblocking network that satisfies the given set of constraints, using linear programming. This makes it possible to compare any candidate network design to a lower bound on the cost of the best possible network. Very often, this allows us to show that the network design produced by a general design algorithm for a specific instance, is much closer to optimal than what is implied by the worst-case analysis for that algorithm. Our current technique yields lower bounds that are usually close to the cost of an optimal network for problem instances with only source/sink capacities and hierarchical clustering constraints [4]. For instances with simple distance constraints, the gap between lower bounds and the cost of designs produced by our best algorithms is typically a factor of four. We believe that in this case, much of the gap is due to weakness in the lower bound, rather than the network design algorithms. One of our key outstanding open problems is the development of better lower bound techniques for problems with distance constraints.

There are two other respects in which our formulation of the network design problem idealizes reality. The first is our use of link costs that increase linearly with capacity. As argued above, this does not lead to any serious inaccuracy when link group capacities are large enough to require multiple parallel links of the highest capacity available, which we believe will be a common case. For networks requiring smaller link group capacities, we

have the following new considerations (1) non-linearity of link costs, relative to capacity, (2) fragmentation of link group capacity over multiple physical links and (3) the question of how to configure a link group of specified capacity using a fixed set of link types so as to provide the required capacity at the lowest cost. The fragmentation question has been addressed in [5] where it is shown how to account for worst-case fragmentation assuming some specified limit on the maximum rate of an individual virtual circuit. The design of the best possible link group, using a fixed set of link types turns out to be equivalent to the well-known and theoretically intractable knapsack problem. However, in practice, the number of alternatives is small enough to allow solution by exhaustive enumeration or a simple dynamic programming algorithm [5].

The second way in which our model idealizes reality is our common assumption that the switch costs can be simply allocated to the links. While, we have argued that this is a reasonable assumption to make, in the context of efficient switching system architectures, it can lead to network designs requiring switching systems with greater capacity than are commercially available. This makes it necessary to replace large switches with a functionally equivalent set of smaller switches. In [5] we show how this can be done, but the cost of these "equivalent switch groups" grows rapidly enough that we can no longer just allocate the cost of a switch to the transmission links without committing a significant error. Accounting for switch costs more explicitly will be important in future network design algorithms.

References

[1] V. E. Beneš. *Mathematical Theory of Connecting Networks and Telephone Traffic.* Academic Press, NY, 1965.

[2] M. Bern and P. Plassman. "The Steiner Problem with Edge Lengths 1 and 2." *Info. Proc. Letters,* 32, 1989, 171–176.

[3] C. Clos. "A Study of Nonblocking Switching Networks." *Bell Systems Technical Journal,* 32, 1953.

[4] , J. A. "Approximation Algorithms for Configuring Nonblocking Communication Networks." Washington Univ. Computer Science Department, Doctoral Dissertation, 5/94.

[5] Fingerhut, J. A., Rob Jackson, Subhash Suri, Jonathan Turner. "Design of Nonblocking ATM Networks." Washington Univ. Computer Science Department, WUCS-9603, 1/96.

[6] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.

[7] A. Gersht and R. Weihmayer. "Joint Optimization of Data Network Design and Facility Selection." *IEEE J. on Selected Areas in Communications*, 12/90.

[8] R. J. Gibbens and F. P. Kelley. "Dynamic Routing in Fully Connected Networks." *IMA J. of Mathematical Control and Information*, 1990.

[9] A. Kershenbaum P. Kermani and G. Grover. "MENTOR: An Algorithm for Mesh Network Topological Optimization and Routing." *IEEE Transactions on Communications*, April 1991.

[10] T. L. Magnanti, P. Mirchandani and R. Vachani. "Modeling and solving capacitated network loading problem." Working paper, Operations Research Center, MIT, 1991.

[11] R. Melen and J. S. Turner. "Nonblocking Multirate Networks." *SIAM J. on Computing*, 4/89.

[12] R. Melen and J. S. Turner. "Nonblocking Multirate Distribution Networks." *IEEE Transactions on Communications*, vol. 41(2), 1993, pp. 362–369.

[13] M. Minoux. "Network Synthesis and Optimum Network Design Problems: Models, Solution Methods and Applications." *Networks*, 1989.

[14] N. Pippenger. "Telephone Switching Networks." *Proc. Symp. on Applied Mathematics*, 1982.

[15] Turner, Jonathan S.. "An Optimal Nonblocking Multicast Virtual Circuit Switch," *Proceedings of Infocom*, 6/94.