Network Global Expectation Model: A Statistical Formalism for Quickly Quantifying Network Needs and Costs

Steven K. Korotky, *Fellow, IEEE, Fellow, OSA*

***Abstract*—In the model presented here, we evaluate expectation values over the entire network to obtain a multimoment description of the required quantities of key network and network element (NE) resources and commensurate network costs. This approach naturally and analytically connects the global (network) and local (NE) views of the communication system and thereby the model can be used as a tool to gain insight and very quickly provide approximate results for the preliminary evaluation and design of dynamic networks. Further, the global expectation model can serve as a valuable guide in the areas of NE feature requirements, costs, sensitivity analyses, scaling performance, comparisons, product definition and application domains, and product and technology roadmapping. We illustrate the application of the techniques to backbone, fiber-optic transport networks.**

***Index Terms*—Communication system economics, communication systems, modeling, networks, optical communication.**

I. Introduction

THE TECHNOLOGY and architecture for circuit and packet communication networks continue to evolve and converge [1]. Fundamental to the comparison and selection of network architectures and their technological implementations is the total cost of ownership of the network. This cost includes the expenses for capital equipment (CAPEX), network operation (OPEX), and network management (MANEX). While operational and management expenses represent the largest share of the total cost of ownership, capital costs are a considerable and highly visible portion of the initial investment. Equipment cost is therefore a very important factor in the choice of architecture and technology. In this paper, we focus our consideration on the cost of the physical gear that constitutes the network and describe and illustrate a model for very quickly gauging network equipment needs and costs.

In the model, network global expectation values are used as a multimoment description of the required quantities of key network and network element (NE) resources and commensurate network costs. This approach naturally, analytically, and accurately connects the global (network) and local (NE) views of the communication system. As a result, the model can be used as a tool to gain insight and quickly provide approximate results for preliminary network evaluation and design, element feature requirements, costs, sensitivity analyses, scaling performance, comparisons, product definition and application domains, and roadmapping. Previously, we have summarized some useful aspects and preliminary results of the model when applied to backbone fiber-optic transport networks [2]. Here, we provide a more complete, general formulation and description, as well as the inclusion of total network costs, more detailed NE cost structures, nonregular networks, and the variances of key variables.

As will become evident, the framework of the model is readily adaptable to various levels of detail and approximation and to a wide range of networks. Our goal regarding the utility and character of the model has been to permit results to be computed very fast with useful accuracy for a very wide range of network sizes and thereby to provide valuable understanding and guidance. To accomplish this, our approach has been to formulate analytic or closed-form relationships among the important input and output network variables via formal derivation, considered approximation, and semi-empirical observation. Our intent has been to complement, not supplant, the more detailed and more accurate, but computationally intensive and very time-consuming, network planning and optimization tools based on numerical simulation [3]–[12]. Such detailed analyses remain critical to the thorough engineering, costing, pricing, and evaluation of specific networks and network products.

The present model can serve as a precursory tool with which to survey the landscape of options, as a means to interpolate and extrapolate the more micro-scale analyses, to investigate scaling performance, and to cross-check complex numerical simulation tools. The analytic nature of the network global expectation model also enables results for arbitrary-size networks to be computed extremely quickly using very modest computational resources and is therefore useful for network analyses in dynamic operating and technological environments, such as encountered in evolving provisionable and survivable backbone networks. We suggest that the uncomplicated and transparent accounting of NEs, systems, and costs inherent in the network global expectation model can constitute a framework for the cooperative exchange of critical planning information on evolving network needs across the many sectors of the communication business. In analogy to the silicon electronics industry, coordination of the materials, components, subsystems, systems, networks, operations, applications, user, and investor communities through more public roadmapping may allow rapid advance in capability and service with higher efficiency and less market volatility than would otherwise be the case.

Our work is divided into several major parts. In the present exposition, we develop the general formalism of the global

Manuscript received July 11, 2003; revised December 2, 2003.
The author is with the Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733-3030 USA (e-mail: skk@lucent.com).
Digital Object Identifier 10.1109/JLT.2004.825756

network expectation model and illustrate its application by considering single-tier backbone networks with location-independent traffic demands. In the future, we will consider the refinement and extension of the approach to a wider set of topologies, architectures, and traffic profiles. While the methodology we present is very general, for specificity throughout this paper, we describe its application in the context of mesh networks.

II. Network Global Expectation Model

A. Costs and Expectation Values

As the cost of the network for a specified set of features is considered the metric for comparison of architectures and technologies, we build the model and begin its description from this perspective. Within the limits of consideration set forth previously, the total network cost is exactly the sum of the costs of the constituent parts, or elements, of the network. This fundamental accounting of costs may be written mathematically as

$$C_T \equiv \sum_i c_i \tag{1}$$

where C_T is the total network cost and c_i is the unit cost of the ith component. (Here and throughout this paper, the symbolic notation Σ indicates the summation over the various contributing terms, in this case the many individual components.)

It is usual that there are many components of a given type used throughout the network, and these identical parts share a common cost. In this case, using the associative, commutative, and distributive properties of the field of real numbers, (1) may be rewritten as

$$C_T = \sum_i \nu_i c_i \tag{2}$$

where again C_T is the total network cost, ν_i is the number of NEs of type i, and c_i is the corresponding unit cost of NE of type i.

Without loss of generality, we may assume that the technology and corresponding unit costs c_i, of the NEs used to construct the network are known, i.e., given *a priori*. The challenge of network design is to determine the number ν_i and placement of each of the NEs of the given types to minimize the total network cost under the constraint to service a specified traffic demand among the network terminations located at specific geographic locations. The strategy of the present model is to carefully estimate the products of the NE counts and respective costs while satisfying the external constraints, and thereby to estimate the total network cost using (2), but without explicitly establishing knowledge of the placement of every individual component.

The sum in (2) does not distinguish among the various categories of NEs but considers each contributing type as atomic, i.e., indivisible. Without changing the value of the sum, we may collect terms that are logically related to one another into a cost subtotal for larger categories of elements. Denoting a general set of categories as $\{\alpha\}$, we may then rewrite (2) as

$$C_T = \sum_\alpha \sum_i \nu_i(\alpha) c(\alpha)_i. \tag{3}$$

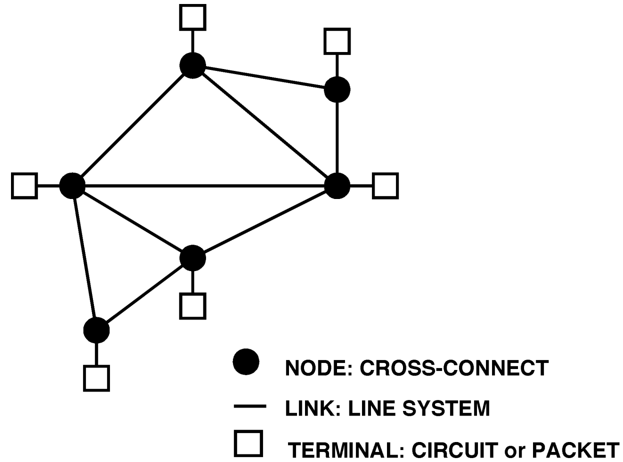


Fig. 1. Mesh network architecture. The average degree of node is $\langle\delta\rangle = 3$ for this network topology of $N = 6$ nodes and $L = 9$ links.

One useful subdivision for separating costs is based on collecting the costs for signal transmission (TRANS) and signal bandwidth management (BWM) into separate terms. In this case, (3) may be arranged into the form

$$C_T = \sum_{\mathrm{TRANS}} \nu_i c_i + \sum_{\mathrm{BWM}} \nu_j c_j. \tag{4}$$

The transmission term might include, for example, objects such as optical transceivers (OTs), optical multiplexers (OMUX), and optical amplifiers (OA). The bandwidth management term might include objects such as add/drop multiplexers (ADMs), Internet protocol routers (IPRs), multiservice platforms (MSP), electronic cross connects (EXCs), optical add/drop multiplexers (OADMs), and optical cross connects (OXCs). Of course, which objects are to be associated with particular categories is a matter of architectural choice.

The abstract representation of the mesh networks we consider is depicted in Fig. 1. The network consists of nodes, where traffic may enter and leave the network; terminals connected to nodes, which are the sources and sinks of the traffic; and links (or edges), which represent the physical segments over which the traffic may be carried, or transported, between nodes. The total number of nodes and links of the network are denoted by N and L, respectively. For concreteness in Fig. 2, we illustrate an example of the mesh networks to which the model may be applied.

As suggested by the view of the networks illustrated in Figs. 1 and 2, the total network cost C_T may also be represented by terms that correspond to the L links and N nodes of the network, viz.

$$C_T = \sum_l^L c_l + \sum_n^N c_n \tag{5}$$

or

$$C_T = \sum_{\mathrm{LINKS}} c_l + \sum_{\mathrm{NODES}} c_n \tag{6}$$

where c_l is the cost of the lth link, and c_n is the cost of the nth node. If we multiply the first term of (5) by L/L and the second

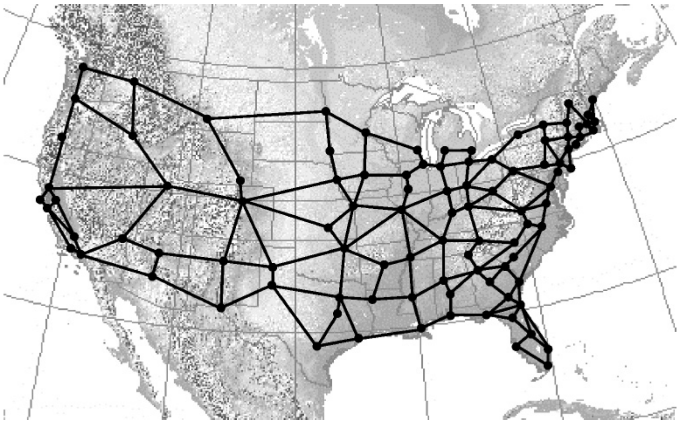


Fig. 2. Prototypical backbone network. Illustrated is a hypothetical core fiber transport network indicative of the larger inter-exchange carriers of the continental United States. This example network consists of 100 nodes and 171 links [6]. The average degree of nodes is $\langle \delta \rangle = 3.4$, and the average number of minimum hops between node pairs is $\langle h \rangle = 6.6$. (*Background relief map courtesy of the U.S. Department of the Interior.*)

term by N/N and note that the expectation value $\langle q \rangle$, or average, of a set of values $\{q_i\} i = 1, m$ is by definition

$$\langle q \rangle \equiv \frac{1}{m} \sum_i^m q_i \tag{7}$$

then we may write (5) as

$$C_T = L\langle c_l \rangle + N\langle c_n \rangle. \tag{8}$$

Thus, as expressed in (8), the exact cost of the network may be considered as the sum of the expectation value of the cost of a link times the number of links and the expectation value of the cost of a node times the number of nodes. The global expectation values $\langle c_l \rangle$ and $\langle c_n \rangle$ are themselves explicitly

$$\langle c_l \rangle = \sum_i \langle v_i \rangle_l c_i \tag{9a}$$

and

$$\langle c_n \rangle = \sum_j \langle v_j \rangle_n c_j. \tag{9b}$$

Note, throughout this paper, we will continue to use the bracket notation $\langle q \rangle$ to denote the expectation value of a variable—in this case, the mean value of q. In instances when the corresponding set $\{q\}$ of an expectation value $\langle q \rangle$ may be ambiguous, we may follow the right bracket of the expectation value by a subscript to provide clarification. For example, in (9a) $\langle \nu_i \rangle_l$ indicates an expectation value over the set of links $\{l\}$, and in (9b) $\langle \nu_j \rangle_n$ indicates an expectation value over the set of nodes $\{n\}$. In addition, regarding expectation values, we point out that here the elements q_i of the set $\{q\}$ define a distribution, rather than represent samples of a variable associated with either a discrete or continuous probability distribution.

At this point, we briefly comment on the nature of the approximations of the network global expectation model. The relationship of network cost to link, node, and terminal costs embodied in (8) and (9) may appear intuitively obvious and, as such, could have served as the starting point of the discussion. However, we have chosen instead to begin using (1) to firmly establish that the use of expectation values, or averages, to determine the total network cost is *not* an approximation but is exact. The approximations of the network global expectation model reside instead in the estimation of the expectation values of the quantities of NEs, $\langle \nu \rangle$. Consequently, the predictive capability of the model will depend upon the accuracy of the estimations of these mean values and the applicability of other related assumptions, such as the demand model. As we shall see, for many variables the expectation values may be computed exactly from the input variables for a given demand model, while for other variables, it is necessary to introduce semi-empirical approximations. In the following sections of this paper we turn next, therefore, to the discussion of a basic model of the key network and NE variables, and costs. Through this basic model, we convey the immediate utility and power of the methodology. In future works, we introduce and illustrate additional refinements of the models of the expectation values to extend further its applicability and accuracy.

B. Network and Primary Model Variables

Referring to Figs. 1 and 2, we define a communication network as the combination of a network *graph* G, consisting of a set of N nodes $\{n_i\}$ and set of L connecting two-way links, or edges $\{l_i\}$, and a network *traffic*. The network graph may be represented by the symmetric matrix $[g]$ with elements g_{ij} [13]. The pairwise two-way communication traffic between terminals located at different nodes may be represented by the symmetric demand matrix $[d]$ with elements d_{ij} and the total ingress/egress traffic T.

The matrix elements g_{ij} are either 0 or 1 in value and specify whether a pair of nodes is connected via a physical link. The summation of all the values of the matrix elements of $[g]$ yields the number of one-way links L_1, which is twice the number of two-way links L_2. The demand matrix elements d_{ij} are either 0 or a positive integer and denote the magnitude of the terminal-to-terminal traffic in quantized units of some basic measure of communication bandwidth, such as a standardized channel bit rate B. The summation of all the values of the matrix elements of $[d]$ yields the number of one-way demands D_1, which is twice the number of two-way demands D_2. Generally, the diagonal elements of $[g]$ and $[d]$ are zero. Note, sometimes the demands are also referred to as logical links to highlight the relationship to and distinction from the physical links (edges).

Often the channel bit rate is not explicitly given for the network of interest. Instead, the total ingress/egress (i.e., total terminal input/out (I/O)) traffic T and number of demands are specified. In that case, a value of the traffic demand bit rate, denoted τ, must be deduced, and from this an appropriate value of B may be chosen. For example, if the total network traffic is 1 Tb/s, and this traffic is the result of 600 unit demands, then the unit demand rate is precisely $\tau = 1$ Tb/s/600 $\cong$ 1.667 Gb/s. This demand rate is less than and about equal to the synchronous optical network (SONET) STS-48 payload rate of ~2.378 Gb/s. Therefore, in this case, an appropriate value of the unit channel rate might be chosen to be the SONET STS-48 signal rate of $B \cong$ 2.488 Mb/s. In this paper, we consider the total two-way

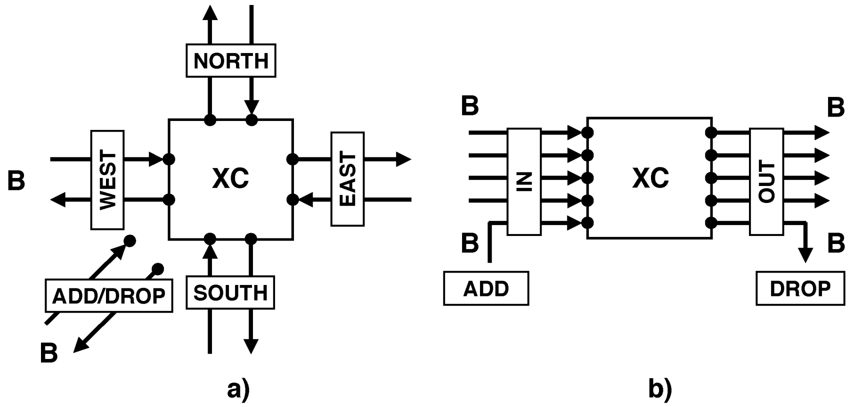


Fig. 3. Cross-connect views and one-way and two-way demands. (a) The cross-connect and line systems are arranged to illustrate five two-way ports (north, south, east, west, and add/drop) appearing on a cross connect. (b) The same cross-connect and line systems are re-arranged to illustrate five one-way ports (five inputs and five outputs). Note that the numbers of one-way and two-way ports are identical, i.e., $P_1 = P_2$. In addition, the channel bit rate B, or alternatively the traffic demand bit- rate τ, describes both the one-way and two-way traffic between terminals, which are indicated here as add/drop.

traffic T, which is one-half the total one-way traffic T_1, to be an independent variable and for τ to be a dependent variable. Having chosen T as an independent variable, we now have a complete set of model inputs, namely

> Primary Model Input Variables: $G\ (N, L)$, D, and T together with a demand model.

As we shall see, all other variables of interest may be determined from these.

In counting quantities such as links, demands, and traffic, it is necessary to distinguish between one-way (simplex) and two-way (duplex) variables. As we have indicated previously, the number of two-way links, demands, and traffic is one-half the corresponding number of one-way values. For future reference, we formally summarize these relationships as

$$\text{Links :} \qquad L \equiv L_2 = \frac{L_1}{2} \tag{10a}$$

$$\text{Total Traffic :} \qquad T \equiv T_2 = \frac{T_1}{2} \tag{10b}$$

$$\text{Total Demands :} \qquad D \equiv D_2 = \frac{D_1}{2}. \tag{10c}$$

Having stated this, we also note that it is usual to define a two-way channel of bandwidth B as the combination of two one-way channels, XY and YX, each of bandwidth B, i.e., the single value B describes both the one-way and two-way channels, which is evident in the example depicted in Fig. 3. In addition, considering the trivial case of two nodes $N = 2$ and one two-way link $L = 1$, the reader will appreciate that the total one-way traffic is $T_1 = 2B$, and the total two-way traffic is $T = T_2 = B$. Of course, so long as one-way or two-way variables are used consistently, or the proper conversion is made, the results and conclusions are the same. For example, $B = T_2/D_2 = T_1/D_1$.

The output variables that are determined by the network global expectation model given the small number of inputs are many. Among them are the traffic demand bit rate and expectation values and variances for the degree of node, number of hops, wavelengths on a link, traffic on a link, restoration capacity, number of ports on a cross connect, total capacity of a cross connect, and percentage add/drop at a node. With these expectation values and a cost model for the individual elements, we can also compute the total network cost. In the next section, we derive expressions for important output variables in a logical sequence using the expectation value formalism. In almost all cases, the expressions we derive are valid independent of the demand model. For those cases where it is necessary to assume a demand model to derive or illustrate an analytic form of the dependencies, in this paper, we consider the case of location-independent demand of which uniform demand and random demand are particular instances. In the future, we plan to extend the analytic model by considering a broader range of topologies, architectures, and demand scenarios.

III. Single-Tier Networks With Location-Independent Demands

A. Expectation Values of NE Quantities

To introduce the global expectation model, we first consider a single-tier network consisting of a set of peer nodes and uniform demand which implies a complete set (fully connected) of equal numbers of demands among all the terminals at the nodes of the network. While this may seem restrictive, in fact the network global expectation model can be applied to a wide range of network topologies, architectures, and demand profiles. This will become evident as we formulate the expectation values and derive general relationships that are independent of the details of the topology, architecture, and demand.

Most core networks carry symmetric traffic between nodes, and so working with two-way variables is the norm. However, in some instances, visualizing and counting one-way variables may be more intuitive, such as tracking a one-way demand from source to destination. Of course following two-way demands from termination to termination is equivalent. In the following exposition, we will explicitly develop expressions using both one-way and two-way input variables for completeness and utmost clarity. The reader will observe that in very many cases, the definition of output variables is such that the values do not change when switching between the one-way and two-way perspectives, such as we have seen in Section II-B for the traffic channel bit rate B.

As an illustrative aide for the reader, throughout this paper, we will apply the model to estimate key characteristics of two example networks. One is the network depicted in Fig. 2, which consists of 100 nodes and 171 links, uniform demand, and total two-way network traffic of 5 Tb/s. The second example network (not shown) has a topology similar to the first—being a mesh network derived from the first example and covering the same geographic area with nearly identical ratio of the number of links to the number of nodes—and again supports uniform demand. This second example differs from the first in that it consists of a smaller number of nodes and links and serves a smaller total traffic. Specifically, it consists of 25 nodes and 42 links, uniform demand, and total two-way traffic of 1 Tb/s.

B. Number of Demands

The number of nodes N, the total two-way traffic T, and the number of two-way links L are inputs of the model. The traffic demand is also an input of the model. The total number of demands is explicitly and, of course, straightforwardly related to the numbers of demands terminating at the individual nodes. We may relate the one-way demands originating at node i to the elements of the demand matrix $[d]$, viz. $d_i = \Sigma^N d_{ij}$. Summing the origininating one-way demands, we may then relate the total one-way and total two-way demands to the mean number of demands originating at a node $\langle d \rangle_n$ as

$$D_1 = \sum_i^N d_i = \frac{N}{N} \sum_i^N d_i = N \langle d \rangle_n \tag{11a}$$

and

$$D \equiv D_2 = \frac{D_1}{2} = \frac{1}{2} N \langle d \rangle_n. \tag{11b}$$

These expressions are independent of the details of the demand model. The uniform demand model specifies that there is a one-way demand from every terminal to every other terminal, or a two-way demand between every terminal–terminal pair of the N nodes. Thus, for uniform demand

$$\langle d \rangle_n = N - 1 \tag{11c}$$

and

$$D_1 = N(N-1) \tag{11d}$$

$$D \equiv D_2 = \frac{N(N-1)}{2}. \tag{11e}$$

The number of two-way demands (logical links) for our example network of $N = 100$ nodes and $L = 171$ physical links is $D = 4950$. The number of two-way demands for our second example network of $N = 25$ nodes and $L = 42$ links is $D = 300$.

C. Traffic Demand Bit-Rate

The value of the traffic demand bit rate τ can be computed exactly as the ratio of the total ingress/egress traffic T and total number of two-way network demands D terminating at all nodes. We have

$$\tau \equiv \frac{T_1}{D_1} = \frac{T_2}{D_2}. \tag{12a}$$

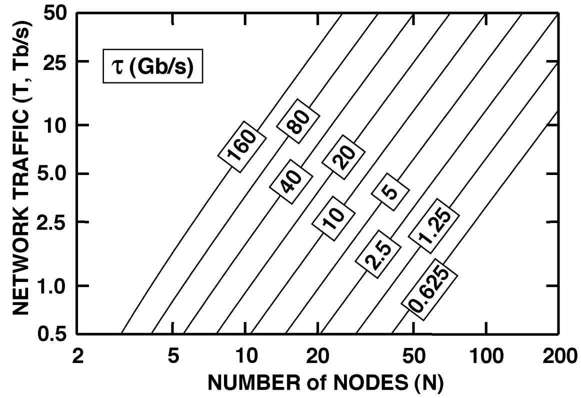


Fig. 4. Traffic demand bit rate. The traffic demand bit rate $\tau(N, T)$ for uniform demand is graphed as a function of the number of nodes N and total two-way traffic T using a contour plot. Contours of constant τ are labeled in units of gigabits per second.

We emphasize that the total traffic T and total number of demands D define the traffic demand bit rate τ, as indicated by the relationship expressed in (12a), which is independent of the demand model. Said differently, as the total traffic and the number of demands define τ, the value of τ is uniquely specified, and as such, its variance is exactly zero. If we specify a demand model, then the particular value of τ for that model may be determined using (12a). In the case of uniform demand, substituting (11e) in (12a) yields

$$\tau \equiv \frac{T}{\left[\frac{N(N-1)}{2}\right]}. \tag{12b}$$

The traffic demand bit rate for uniform demand is plotted as a function of the number of nodes N and total network traffic T in Fig. 4.

The traffic demand bit rate for our example network of $N = 100$ nodes, $L = 171$ links, and total traffic of $T = 5$ Tb/s is $\tau = 1.01$ Gb/s. This may be compared with $\tau = 3.3$ Gb/s for our example network of $N = 25$ nodes, $L = 42$ links, and total traffic of $T = 1$ Tb/s. The channel bit rate is smaller for the larger network because the number of demands for the larger network is significantly greater than for the smaller network.

D. Degree of Node

1) Mean Value: The average degree of a node $\langle \delta \rangle$, i.e., $\langle \delta \rangle_n$, is calculated straightforwardly by summing the number of one-way (directed) links and by dividing by the number of nodes. Referring to the matrix representation $[g]$ of the network graph, we have

$$\delta_i = \sum_j^N g_{ij} \tag{13a}$$

and so

$$\langle \delta \rangle = \frac{1}{N} \sum_i^N \sum_j^N g_{ij} = \frac{L_1}{N} = \frac{2L_2}{N} = \frac{2L}{N}. \tag{13b}$$

This compact expression for $\langle \delta \rangle$ is exact and independent of the demand model.

2) Variance and Standard Deviation: The variance $\sigma^2(q)$ and standard deviation $\sigma(q)$ of the set of values for the network variable q are defined by [14], [15]

$$\sigma^2(q) \equiv \frac{1}{m}\sum_{i}^{m}(q_i - \langle q \rangle)^2 \tag{13c}$$

which may be rewritten as

$$\sigma^2(q) \equiv \langle q^2 \rangle - \langle q \rangle^2. \tag{13d}$$

Again, we note that the set $\{q\}$ is not a sampled data set but defines the distribution. We also alert the reader that the standard deviation of a network variable is not an indication of the accuracy or error of the model, but rather it is a measure of the variation of the number of NEs or subsystems from locale to locale across the network. Note too that, in general, the mean and variance are independent variables. Thus, for example, the total cost for bandwidth management may be accurately predicted even while some nodes are smaller and cost less, and others are larger and cost more.

The variance of the degrees of nodes is by definition

$$\sigma^2(\delta) \equiv \langle \delta^2 \rangle - \langle \delta \rangle^2 \tag{13e}$$

and therefore, like δ_i and $\langle \delta \rangle$, $\sigma^2(\delta)$ is a function only of the network graph G. Note, however, unlike $\langle \delta \rangle$, there is no closed-form expression for $\sigma^2(\delta)$ as a function only of N and L. Rather, the variance of the degrees of nodes implicitly depends upon the details of the network connectivity and must be computed from a representation of the graph, such as $[g]$ or an equivalent link list. If the network graph, or equivalently the link list, is provided then functions of the degrees of nodes, such as the variance, may be computed exactly.

Note, as $\langle \delta \rangle$ and L are directly proportional, and the variance of δ is more closely related to $[g]$, in some situations, it may be useful to consider $\langle \delta \rangle$ as the independent input variable and L as the dependent output variable.

For our example network of $N = 100$ nodes and $L = 171$ links, the mean degree of node is $\langle \delta \rangle = 3.4$. The standard deviation of the nodal degree obtained from the network graph (Fig. 2) is $\sigma(\delta) = 1.1$. By design, the mean degree of node and standard deviation of the nodal degree for the second example network of $N = 25$ nodes and $L = 42$ links are also $\langle \delta \rangle = 3.4$ and $\sigma(\delta) = 1.1$.

E. Number of Hops

1) Mean Value: The number of hops between a pair of terminals is defined as the minimum number of links traversed by a demand between the terminating node pair. Algorithms for determining the minimum number of hops h_{ij} between node pairs (i,j) from the matrix representing the network graph $[g]$ are well known, and so $[h]$ and $\langle h \rangle$ may be readily computed given a demand model [13]. The expectation value of the minimum number of hops is over the set of demands, i.e., $\langle h \rangle_d$, and formally we have

$$\langle h \rangle = \frac{1}{D}\sum_{i<j}^{D} h_{ij} = \frac{1}{2D}\sum_{i,j}^{D} h_{ij}. \tag{14a}$$

If the network graph and demands are provided, then we may proceed to compute $\langle h \rangle$ exactly. However, we may also approximate $\langle h \rangle$ for location-independent (uniform or random) demands with knowledge only of the number of nodes and number of links, as discussed subsequently.

The dependency of the average number of hops on the number of nodes N and number of links L may be formulated by considering the schematic of the network graph. If we visualize the outer boundary of the N nodes of a planar network arranged roughly as a square with $\sqrt{N}$ nodes on each of the two orthogonal sides, we appreciate that the characteristic distance scale of the network measured in units of hops scales as $\sqrt{N}$ for uniform demand. We also realize that the mean number of hops decreases as the number of links L increases for fixed N. An approximate analytic relationship describing the dependency of the mean number of hops on the number of nodes N and the mean degree of the nodes $\langle \delta \rangle$ may be derived by considering a single node at the center of a regular mesh network of constant degree $\delta > 2$ [2]. By dividing the network into δ sectors centered on the selected node and computing the mean number of hops from the selected node to the $(N-1)/\langle \delta \rangle$ nodes within the sector using continuous integration to replace the discrete summation, we find that the mean number of hops is approximately $\langle h \rangle \cong 0.94\sqrt{[(N-1)/\langle \delta \rangle]}$. This expression slightly under-predicts the correct result in the special case where each node is connected directly to every other node via a dedicated physical link, i.e., $\delta = N-1$ and $\langle h \rangle \equiv 1$. Brute force evaluation of the mean number of hops for regular networks of constant degree for $\delta = 3$ and $\delta = 4$ except for the nodes at the perimeter yields $\langle h \rangle \cong 1.2\sqrt{[N/\langle \delta \rangle]}$, which slightly over-predicts the mean number of hops for the special case of $\delta = N-1$ and $\langle h \rangle \equiv 1$.

In the spirit of providing accurate compact analytic expressions for all variables for a wide range of networks, we have analyzed the average number of hops for several prototypical networks that were designed to be survivable under all possible single link failures. (Note, the failure of a single link implies the simultaneous failure of all demands appearing on the specified edge, which may be a very large number of interterminal demands). This feature of network survivability translates into the requirement that the degrees of nodes for all nodes be greater than or equal to two, i.e., $\delta \geq 2$. The exact results for the mean number of hops were fitted using the method of least squares deviation to determine the single coefficient of proportionality that best describes the data for all the networks considered. In total, data for 14 mesh networks with numbers of nodes spanning the range $4 \leq N \leq 100$ and average degree of node spanning the range $2.5 \leq \langle \delta \rangle \leq 5$ were included [2], [7], [10]. We find that the expectation value of the number of hops for these networks with uniform demand may be expressed semi-empirically by the relation

$$\langle h \rangle \cong 1.12\sqrt{\frac{N}{\langle \delta \rangle}} \tag{14b}$$

with a standard deviation of approximately 10%, and more accurately by the semi-empirical relation

$$\langle h \rangle \cong \sqrt{\frac{(N-2)}{(\langle \delta \rangle - 1)}} \tag{14c}$$

with a standard deviation of approximately 5%.

We remind the reader that these approximate formulas may be applied to the case of uniform or random demand, and for fixed network topology clearly we expect the average number of hops to decrease for distance-dependent demand models that weigh shorter distance demands more heavily than longer distance demands. Note also that for nominally linear networks, such as a ring ($\delta = 2$), the number of hops scales as NN.

The estimate of the mean number of hops for our example network of $N = 100$ nodes and $L = 171$ links using (14c) is $\langle h \rangle \cong 6.1$, which may be compared with the actual mean of $\langle h \rangle = 6.6$. For the example network consisting of $N = 25$ nodes and $L = 42$ links, the mean number of hops using (14c) is approximately $\langle h \rangle \cong 3.0$.

2) Variance and Standard Deviation: The variance of the number of hops may be computed from $[h]$ using (13); however, we have not found a need to compute $\sigma^2(h)$ explicitly for the analyses that follow. We note that the range of hops extends from 1 to some maximum number H, which is often referred to as the diameter of the network.

F. Demands on Link

1) Mean Value: It is evident that as a demand d_{ij} is routed across the network between terminating nodes (i, j) that the demand occupies a unit of transmission capacity on each of the links connecting the nodes. The minimum number of links occupied by a demand is, of course, the minimum number of hops h_{ij} from node i to node j. Consequently, the average number of demands carried on a link in the absence of extra capacity for restoration is

$$\langle W^o \rangle = \frac{1}{L}\sum_{l}^{L} D_l \cdot 1 = \frac{1}{L}\sum_{i,j}^{D} 1 \cdot h_{ij} = \frac{1}{L}\frac{D}{D}\sum_{i,j}^{D} h_{ij} = \frac{D\langle h \rangle_D}{L} \tag{15a}$$

which may rewritten in the convenient form

$$\langle W^o \rangle = \frac{\langle d \rangle \langle h \rangle}{\langle \delta \rangle} \tag{15b}$$

using (11b) and (13b). This important new result is exact and valid independent of the demand model; however, the value of $\langle h \rangle$ is implicitly dependent upon the demand model, as discussed previously. In the cases of uniform or random demand, if an approximation for $\langle h \rangle$ such as (14b) or (14c) is used to compute $\langle W^o \rangle$, then of course the result is also approximate, and the relative error of $\langle h \rangle$ determines the relative error of $\langle W^o \rangle$.

For uniform demand, we may substitute for $\langle d \rangle$ using (11c) in (15b) to obtain

$$\langle W^o \rangle = (N-1)\frac{\langle h \rangle}{\langle \delta \rangle}. \tag{15c}$$

Using (15c), the mean number of channels carried on a link for the first example network of $N = 100$ nodes and $L = 171$ links ($\langle \delta \rangle = 3.4$, and $\langle h \rangle \cong 6.1$) is estimated to be $\langle W^o \rangle \cong 178$. Similarly, the mean number of channels on a link for the second example network of $N = 25$ nodes and $L = 42$ links ($\langle \delta \rangle = 3.4$, and $\langle h \rangle \cong 3.0$) is estimated to be $\langle W^o \rangle \cong 22$.

2) Variance and Standard Deviation: As suggested by the dependencies on $\langle d \rangle$, $\langle \delta \rangle$, and $\langle h \rangle$ expressed in (15b), variations in the number of channels carried on the individual links of the network may arise from differences in the number of demands terminating at the nodes connected to the links, the degrees of the nodes connected to the link, and also the routing constraints and algorithms. Here, we consider the case of uniform demand and first consider the fluctuations that may arise when the demands are routed across the network under the constraint of minimum hop routing. We observe that, in general, for any pair of terminals, there will be one or more routes of minimum number of hops between the nodes. Consequently, the variation in the number of channels carried on a link will depend upon the selection criteria for choosing from among the set of minimum hop routes, which we refer to as hop-degenerate routes. If we assume that the path is selected at random from the hop-degenerate routes and the probabilities of selecting a link from among the hop-degenerate routes are equal, then we may estimate a reference variance using statistical methods. In particular, for the scenario we have just described, the distribution of the demands among the minimum hop routes may be described by the binomial distribution [16]. We then derive an approximate reference expression for the variance of W^o considering random routing over paths of equal numbers of hops.

Referring to (15a)–(15c) and using (13b), we may write the mean value for the number of channels on a link for uniform two-way demand explicitly as

$$\langle W^o \rangle = \frac{1}{L}\frac{1}{2}\sum_{i}^{N}\sum_{j}^{N-1} h_{ij} = \frac{N(N-1)\langle h \rangle}{2L}. \tag{15d}$$

For a given node pair (i,j), we now consider all the paths of minimum hops h_{ij} between them, and let l_{ij} denote the total number of distinct links among the set of hop-degenerate routes. We label these distinct links using the subscript k and let p_k denote the probability that a link is selected. By construction, the set of probabilities $\{p_k\}$ satisfies

$$h_{ij} = \sum_{k}^{l_{ij}} p_k \tag{15e}$$

and consequently, $p_k \cong h_{ij}/l_{ij}$. As an example, consider an illustrative case when there are three $(r = 3)$ link-disjoint routes of four $(h = 4)$ (minimum) hops between a pair of nodes. In this case, $l_{ij} = r \times h = 3 \times 4 = 12$. As the paths are assumed to be disjoint, we may use (15e) to solve for p_k with the result $p_k = h_{ij}/l_{ij} = h/(rh) = 1/r = 1/3$ for each link.

Substituting (15e) into (15d), we then may write

$$\langle W^o \rangle = \frac{1}{2L}\sum_{i}^{N}\sum_{j}^{N-1}\sum_{k}^{l_{ij}} p_k. \tag{15f}$$

Using the properties of the binomial distribution, the corresponding variance $\sigma^2(W_o)$ is

$$\sigma^2(W^o) = \frac{1}{2L}\sum_{i}^{N}\sum_{j}^{N-1}\sum_{k}^{l_{ij}} p_k(1-p_k) \tag{15g}$$

which may be rewritten as

$$\sigma^2(W^o) = \langle W^o \rangle \left[1 - \frac{1}{N(N-1)\langle h \rangle}\sum_{i}^{N}\sum_{j}^{N-1}\sum_{k}^{l_{ij}} p_k^2\right] \tag{15h}$$

using (15f) and (15e).

To evaluate the sums, we next group the sum over the $N-1$ nodes into sets of constant numbers of hops h. Let there be N_h nodes of h hops, and label each node by the index n. For each node, the number of distinct links among the possible routes of h hops is denoted $l_{n,h}$. If H is the largest value of the set of minimum number of hops, then (15h) for the variance may be written as

$$\sigma^2(W^o) = \langle W^o\rangle \left[1 - \frac{1}{\langle h\rangle}\frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_h^H \sum_n^{N_h}\sum_k^{l_{nh}} p_k^2\right]. \tag{15i}$$

The expression (15h) is exact under the assumption of uniform demand and random routing.

To carry this result further, we next derive an approximation for a planar network of average degree $\langle\delta\rangle$. In this case, the maximum number of hops H satisfies

$$N-1 = \langle\delta\rangle \frac{[H(H+1)]}{2} \tag{15j}$$

and the value of H is related to $\langle h\rangle$ by $H \cong \sqrt{2}\langle h\rangle$. We focus on a single node within the network. The nodes that may be reached in h minimum hops can be identified, and they are approximately $\langle\delta\rangle h$ in number. We next consider the options for routing from the node under consideration to each of the other nodes h minimum hops away. There is a least one possible route, and we denote the number of hop-degenerate routes as r. Next, we identify and count the number of distinct links $l_{n,h}$ among these r hop-degenerate routes. We observe that for the planar network, the number of distinct links $l_{n,h}$ is less than h^2, the latter being the number in the situation when the hop-degenerate routes are link-disjoint paths. Consequently, the probability any one link is selected when choosing a path randomly from among the hop-degenerate routes of the network is greater than $1/h$, i.e.,

$$p_k \geq \frac{1}{h}. \tag{15k}$$

If we assume that all links among the hop-degenerate paths are selected with equal probability, which is not necessarily the case, this expression for the probability a link is selected permits us to formally bound the variance of the number of channels. Substituting (15k) in (15i), carrying out the sums, and using (15j) yields

$$\sigma^2(W^o) \leq \langle W^o\rangle \left[1 - \frac{1}{\langle h\rangle}\right] \tag{15l}$$

and

$$\frac{\sigma(W^o)}{\langle W^o\rangle} \leq \frac{\sqrt{1-\frac{1}{\langle h\rangle}}}{\sqrt{\langle W^o\rangle}} \leq \frac{1}{\sqrt{\langle W^o\rangle}}. \tag{15m}$$

The form of the variance in (15l) is that of a binomial distribution with probability $1/\langle h\rangle$. Thus, we formally approximate the actual distribution by the corresponding binomial distribution $F(W=w)$, which is given by

$$F(W=w) = (w_{\max}|w)p^w(1-p)^{w_{\max}-w}, \quad w=0,1,\ldots,w_{\max} \tag{15n}$$

with

$$p = \frac{1}{\langle h\rangle} \tag{15o}$$

$$w_{\max} \equiv \langle W^o\rangle\langle h\rangle \tag{15p}$$

and

$$(w_{\max}|w) = \frac{w_{\max}!}{[w!(w_{\max}-w)!]}. \tag{15q}$$

The binomial tail probability $F(W \geq w)$ may be determined using the incomplete beta function.

Using (15l), the standard deviation of the number of channels on a link for our example network of $N = 100$ nodes and $L = 171$ links ($\langle\delta\rangle = 3.4$ and $\langle h\rangle \cong 6.1$) is estimated to be $\sigma(W^o) \leq 12$. Recall the mean number of channels on a link was estimated to be $\langle W^o\rangle \cong 178$ for this network. Again using (15l), the standard deviation of the number of channels on a link for our second example network of $N = 25$ nodes and $L = 42$ links ($\langle\delta\rangle = 3.4$ and $\langle h\rangle \cong 3.0$) is estimated to be $\sigma(W^o) \leq 3.8$. The mean number of channels on a link was estimated to be $\langle W^o\rangle \cong 22$ in this case.

In the above consideration of the variation of W^o, we have recognized that usually when traffic is routed and the network is optimized, paths are selected based on criteria such as the minimum number of hops, the shortest distance, or more generally the minimum cost. However, routing solutions that may be proven to be optimal are possible only for relatively small networks, and therefore, additional heuristic constraints are often imposed as strategies to ensure low cost. To minimize the cost of survivable networks, for example, algorithms to balance the traffic among the links are often introduced [7], [8], [10]. By its definition, load balancing deliberately seeks to dampen the variation of the number of channels carried on a link. Clearly, if load balancing is effective then the selection of paths from among the hop-degenerate routes is not random, and $\sigma(W^o)$ should be reduced relative to the value specified by (15l). As a corollary, the ratio of the achieved variance to the value obtained for random routing may be considered a measure of the success of the load-balancing algorithm.

The variance of the number of channels carried on a link derived above is a network global expectation based on routing decisions. We may also consider a local view of the variations and the number of channels carried on a particular link (i,j) and their relationship to the terminating traffic and degrees of the local nodes. We postulate a form for W_{ij} based on (15b) and an heuristic argument based upon the routed traffic. We begin by noting that (15b) may be written to identify the local traffic terminating at the nodes connected to the link (both ends) and the through traffic that passes by both nodes, viz.

$$\langle W^o\rangle = \frac{2\langle d\rangle}{\langle\delta\rangle} + \langle d\rangle\frac{(\langle h\rangle - 2)}{\langle\delta\rangle}. \tag{15r}$$

We observe that the first term corresponds to the division of the terminating traffic among the various links connected to the terminating nodes. Assuming minimum hop routing, to a good approximation, the terminating traffic is equally distributed among all the links connected to the node. This implies a direct correlation of the first term of (15r) to the local degrees of nodes connected to the link. The second term, however, corresponds to the many channels traversing the link that have destinations distributed across the entire network. For the moment, we consider that the traffic is routed to minimize the number of hops, but otherwise no preference among the individual links is imposed. Under these circumstances, we hypothesize that the second term has negligible correlation to the local degrees of nodes and is best described by a combination of the mean value and variations randomly distributed across the network. Therefore, we write the number of channels on a link (i, j) formally as

$$W_{ij} = W_{B/E} + W_{B/T} \tag{15s}$$

with

$$W_{B/E} \equiv d_i \left(\frac{1}{\delta_i} + \frac{1}{\delta_j} \right) - 1. \tag{15t}$$

(The "-1" on the right-hand side in (15t) ensures the proper accounting of the demand between the terminals of node i and node j.) The variable $W_{B/T}$ includes random variations in the number of through channels and satisfies

$$\langle W_{B/T} \rangle \equiv \langle d \rangle \frac{(\langle h \rangle - 2)}{\langle \delta \rangle} + 1. \tag{15u}$$

The variance of $W_{B/T}$ may be estimated using the statistical formalism described previously in (15l) with $W_{B/T}$ replacing W^o, and $\langle W_{\mathrm{B/T}} \rangle$ replacing $\langle W^o \rangle$.

It can be verified by direct computation that the expectation value of $W_{i,j}$ (15s)–(15u) yields $\langle W^o \rangle$ (15r) in the case of location-independent demand, as required. As the second term of (15r) is locally uncorrelated with the first by our hypothesis, the variance of W^o may therefore be expressed as

$$\sigma^2(W^o) \cong \left(\frac{2}{\langle \delta \rangle} \right)^2 \sigma^2(d) + \langle d \rangle^2 \sigma^2 \left(\frac{1}{\delta} \right) + \sigma^2 \left(W_{B/T} \right). \tag{15v}$$

We have already estimated the variance associated with routing decisions implicitly assuming no variation in δ using (15l). Now we may also estimate the relative size of the variance in W^o attributable to variations in the degrees of the nodes. The variations correlated to the local degrees of nodes, i.e., the second term of (15v), can be computed directly from the network graph. For the present we note that for uniform demand $\sigma^2(d) \equiv 0$ and

$$\frac{\sigma \left(W_{B/E} \right)}{\langle W_{B/E} \rangle} \cong \sqrt{\frac{\left[\langle \delta \rangle_n \left\langle \frac{1}{\delta} \right\rangle_n - 1 \right]}{2}}. \tag{15w}$$

Using (15t) and (15w), the mean and standard deviation of the number of A/D channels terminating at the two ends of a link are estimated to be $\langle W_{B/E} \rangle \cong 58$ and $\sigma(W_{B/E}) \leq 13$, respectively, for the larger example network of $N = 100$ nodes and $L = 171$ links ($\langle \delta \rangle = 3.4$, $\langle h \rangle \cong 6.1$, $\langle 1/\delta \rangle = 0.32$). The mean number of channels not terminating at either end of a link is approximately $\langle W_{B/T} \rangle \cong 120$ for this network. For the smaller example network of $N = 25$ nodes and $L = 42$ links ($\langle \delta \rangle = 3.4$, $\langle h \rangle \cong 3.0$, $\langle 1/\delta \rangle = 0.32$), the mean and standard deviation of the number of A/D channels terminating at the two ends of a link are estimated to be $\langle W_{B/E} \rangle \cong 14$ and $\sigma \langle W_{B/E} \rangle \leq 2.8$ using (15t) and (15w). The mean number of channels not terminating at either end of a link is approximately $\langle W_{B/T} \rangle \cong 7.5$ for this example. As required and expected, for both examples, the values of $\langle W_{B/E} \rangle$ and $\langle W_{B/T} \rangle$ sum to the respective value of $\langle W^o \rangle$.

If the terminating demands are not uniformly distributed but instead randomly distributed, then the first term in (15v) proportional to $\sigma^2(d)$, i.e., $\sigma_d^2(W^o)$, also contributes to the variance of W^o, and we have

$$\sigma_d \frac{(W^o)}{\langle W^o \rangle} = \left[\frac{2}{\langle h \rangle} \right] \left[\frac{\langle \sigma(d) \rangle}{\langle d \rangle} \right]. \tag{15x}$$

To close this section, we remind the reader that the expressions for $\langle W^o \rangle$ ((15b) and (15c)) are exact and independent of our estimations of $\sigma(W^o)$.

G. Restoration Capacity

1) Mean Value: The additional capacity added to links to ensure mesh network survivability depends upon the types of failures considered, the restoration strategy, and the blocking characteristics of the cross connects used to redirect the affected traffic over alternate routes. For the purpose of architectural comparisons, network survivability is very often defined in relation to single link failures, i.e., the network is designed and minimally sufficient capacity is deployed to ensure the network can support the traffic and is survivable against all single link failures. As explained previously, this implies the network has sufficient extra capacity to restore all of the simultaneously failed demands sharing the common failed link. Extra capacity is counted in units of additional channel links and is most often reported as a fractional increase above the total number of channel links. Using that convention, the average number of channels on a link including extra capacity for restoration is written as

$$\langle W^\kappa \rangle \equiv \langle W^o \rangle \left(1 + \langle \kappa \rangle \right). \tag{16a}$$

Here, we have introduced the superscript designation κ to W to remind ourselves that the expression accounts for extra capacity for restoration. This expression is independent of the demand model. In considering the individual failure of all the $\delta i + \delta j - 1$ links that are connected to the two nodes at the ends of link (i, j), we model the number of channels on an individual link (i, j), including the extra capacity for restoration as

$$W^\kappa{}_{ij} = W_{ij} + \langle W^o \rangle \kappa_{ij} \tag{16b}$$

where W_{ij} and $\langle W^o \rangle$ are given by (15t)–(15v) and (15s), respectively. The mean value of this model for $W^\kappa{}_{ij}$ yields (16a), as required. Hereafter, we develop formulas for $\langle \kappa \rangle$ and κ_{ij} as functions of the input network variables.

Precisely determining the amount of additional capacity requires a detailed network analysis and is a nontrivial exercise for large mesh networks. Obtaining exact results for general mesh networks when the number of nodes is more than about 20 is presently not practical because of the magnitude and duration of the numerical computations. Thus, some form of heuristic algorithm for routing traffic and assigning restoration capacity is usually employed for large networks [3]–[12].

In considering the extra capacity that must be deployed to ensure survivability against single link failures, a general inverse dependency upon the degree of the nodes is readily recognized and explained qualitatively [17]. For example, for uniform demand a ring network—which by definition has an average degree of node equal to 2—with dedicated protection requires 100% extra capacity relative to the minimum capacity necessary to carry the traffic demand. For this reason, a qualitative relationship between the fractional increase in capacity on a link and the degree of the node to which the link is connected is

$$\kappa \sim \frac{1}{(\delta - 1)}. \tag{17a}$$

However, a strict interpretation of (17a) as an equality can under-predict by one third or more the necessary extra capacity for planar mesh networks and unit, uniform demand when $\langle\delta\rangle$ is greater than 2. To assess the feasibility of using an analytic equation to model the extra capacity, we have fitted the extra capacity determined by detailed calculation and simulation of mesh networks with uniform demands for the case of strictly nonblocking cross connects using the expression

$$\langle\kappa\rangle = \frac{(a-b)}{(\langle\delta\rangle - b)} \tag{17b}$$

where a and b were parameters to be determined semi-empirically.

We considered the results for the extra capacity for path-disjoint shared mesh restoration using a heuristic favoring a small differential path length between working and restoration paths for eight mesh networks [10] and also imposed the condition that $\langle\kappa\rangle = 1$ for $\langle\delta\rangle = 2$. The mesh networks had numbers of nodes N in the range of $4 \le N \le 100$, average degree of node in the range of $2.5 \le \langle\delta\rangle \le 4.5$ and required an average extra capacity in the range of $0.4 \le \langle\kappa\rangle \le 0.9$ [7], [10]. The constraint to describe the ring network exactly using (17b) requires $a = 2$. The best value of b was then determined to be $b = -0.4$. Within the accuracy ($\sigma \cong \pm 17\%$) of the fitted results, the functional form for the extra capacity can be considered to be

$$\langle\kappa\rangle \cong \frac{2}{\langle\delta\rangle}. \tag{17c}$$

For completeness, we note one expression for the extra capacity on the individual links that results in the expectation value of the extra capacity given by (17c) is

$$\kappa_{ij} = \frac{1}{2}\left[\frac{2}{\delta_i} + \frac{2}{\delta_j}\right] \tag{17d}$$

and

$$\begin{aligned}\langle\kappa\rangle &\equiv \frac{1}{L}\sum_{i,j}^{L}\kappa_{ij} \cong \frac{1}{L}\sum_{i,j}^{L}\left(\frac{1}{\delta_i} + \frac{1}{\delta_j}\right)\\ &= \frac{1}{L}\sum_{n}^{N}\sum_{k}^{\delta_n}\frac{1}{\delta_n} = \frac{N}{L} = \frac{2}{\langle\delta\rangle}\end{aligned} \tag{17e}$$

or more explicitly $\langle\kappa\rangle_l = 2/\langle\delta\rangle_n$. (As an aside based on (17e), we note the property that $\langle 1/\delta\rangle_l = 1/\langle\delta\rangle_n$. However, in general, $\langle 1/\delta\rangle_n \neq 1/\langle\delta\rangle_n$, except for regular networks of constant degree δ or as an approximation.)

A slightly more accurate semi-empirical representation ($\sigma \cong \pm 12\%$) of the values of the extra capacities of the networks considered was found to be

$$\langle\kappa\rangle_l = \left\langle\frac{2}{\delta}\right\rangle_n \tag{17f}$$

for which the corresponding local extra capacity is

$$\kappa_{ij} = \frac{1}{2}\frac{\left[\left(\frac{2}{\delta_i}\right)^2 + \left(\frac{2}{\delta_j}\right)^2\right]}{\left[\frac{2}{\langle\delta\rangle}\right]}. \tag{17g}$$

In both cases, it is clear there is a strong correlation between the efficient use of spare capacity for survivability and the degrees of the nodes. Finally, we point out that the additional capacity required for dynamic networks, such as for provisional and/or survivable networks, will be larger if the cross connects are not strictly nonblocking. For example, in the case of wavelength-division-multiplexed line systems and cross connects without wavelength interchange except at the terminations, the increase of the extra capacity for restoration above the minimum value for strictly nonblocking cross connects is typically in the range of only 5–20%, although the management complexity is greatly increased [5]–[10].

For the example network of $N = 100$ nodes and $L = 171$ links ($\langle\delta\rangle = 3.4$), the mean value of the extra capacity to ensure survivability under single link failures is estimated to be $\langle\kappa\rangle \cong 0.58$. As the mean degree of node for the second example network of $N = 25$ nodes and $L = 42$ links is nearly identical to that of the larger network by design, $\langle\delta\rangle \cong 3.4$, the estimate for the mean value of the extra capacity to ensure survivability under single link failures is also nearly the same at $\langle\kappa\rangle \cong 0.60$.

2) Variance and Standard Deviation: In the previous sections, we have modeled the extra capacity on individual links in a manner that is both intuitive and consistent with empirical observations of the total extra capacity. The model for $\{\kappa\}$ depends only upon the degrees of the nodes $\{\delta\}$, and consequently, it is a function of the input network graph G, as stated explicitly in (13a). The variance of κ, $\sigma^2(\kappa)$, can now also be computed straightforwardly using the definition of the variance (13) and the relations for κ (17). For example, for κ_{ij} defined by (17d), we have

$$\frac{\sigma(\kappa)}{\langle\kappa\rangle_l} \cong \sqrt{\frac{\left[\langle\delta\rangle_n\left\langle\frac{1}{\delta}\right\rangle_n - 1\right]}{2}}. \tag{17h}$$

Note also that the deployment of restoration capacity can have the tendency to equalize the capacity allocation on the links, as links with larger working capacity and smaller restoration capacity are able to restore traffic carried on links with smaller working capacity and larger restoration capacity, and vice versa.

H. Traffic on Link

1) Mean Value: The average traffic carried on a link $\langle\beta\rangle$ is the product of the average number of demands on a link $\langle W\rangle$ and the bit rate per demand τ, i.e.,

$$\langle\beta\rangle \equiv \langle W\rangle\tau = \tau\langle h\rangle\frac{D}{L} = \langle h\rangle T/L. \tag{18a}$$

This direct proportionality is independent of the demand model. In Fig. 5, we plot the mean traffic on a link including idle restoration channels for uniform demand as a function of the number of nodes N and total network traffic T.

For our example network of $N = 100$ nodes, $L = 171$ links, and $T = 5$ Tb/s, the mean value of the traffic carried on a link including extra capacity for restoration is $\langle \beta^{\kappa} \rangle \cong 284$ Gb/s. In comparison, the mean value of the traffic carried on a link including extra capacity for restoration for the smaller example network of $N = 25$ nodes, $L = 42$ links, and $T = 1$ Tb/s is $\langle \beta^{\kappa} \rangle \cong 116$ Gb/s.

2) Variance and Standard Deviation: Based on the preceding sections, the variance of β is determined by the variance of W with the result that

$$\frac{\sigma(\beta)}{\langle \beta \rangle} = \frac{\sigma(W)}{\langle W \rangle}. \tag{18b}$$

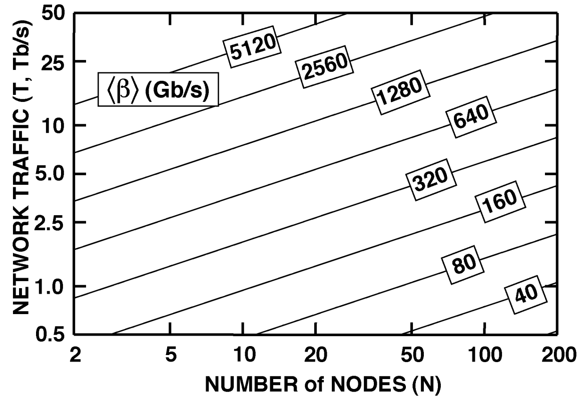


Fig. 5. Mean traffic on link. The mean traffic on a link $\langle \beta^{\kappa}(N, T) \rangle$ for uniform demand with restoration is graphed as a function of the number of nodes N and total two-way traffic T under the constraint $\langle \delta \rangle = 3.5$ using a contour plot.

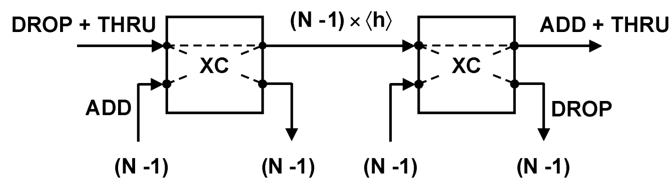


Fig. 6. Demands and cross-connect ports. The figure serves as a guide to count the number of cross-connect ports occupied by a demand as it traverses the network. The relationship among the local add, drop, and through channels is also depicted. Here, the numbers of add and drop demands, each $N - 1$, specifically correspond to the uniform demand model.

I. Number of Ports and Capacity of a Cross Connect

1) Mean Value: Among the key attributes of cross connects are the port count P and total capacity χ. The average number of ports on a cross connect in a mesh network can be determined by counting the number of ports that each demand occupies as it traverses the network, tallying the number of ports for all demands and then dividing by the number of cross connects. By design, a cross connect—of which an ADM is considered a special case—is placed at each node of the backbone network to manage transport bandwidth, and so the number of cross connects is given by the number of nodes N.

Note that, as illustrated in Fig. 3, the number of output ports is usually equal to the number of inputs. In addition, a $P \times P$ cross -connect, which has P inputs and P outputs (or P I/O ports), supports connections among P two-way channels. We first calculate the average number of one-way input ports P_1. Referring to Fig. 6, consider a directed demand that enters, or is added to, the network via the cross connect of the node on the left. Adding the demand requires one input port. Eventually, this demand exits the network. Dropping from the network is accomplished by entering and exiting the cross connect at the destination node, which may be considered the node on the right of Fig. 6. Thus, dropping the demand also requires one input port. In addition, in traversing the network, the demand under consideration occupies input ports at the cross connects of the intervening nodes. Having defined h as the number of internodal hops, the number of intervening cross connects that the demand enters is $h - 1$. Consequently, a one-way demand occupies

$$P_{ij} = 1 + 1 + (h_{ij} - 1) = 1 + h_{ij} \tag{19a}$$

input ports. The total number of input ports occupied by all demands is therefore

$$\begin{aligned} P_t &= \sum_{i,j}^{D_1} [1 + h_{ij}] = \frac{D_1}{D_1} \sum_{i,j}^{D_1} [1 + h_{ij}] \\ &= D_1 \langle 1 + h_{ij} \rangle = N \langle d \rangle [1 + \langle h \rangle] \end{aligned} \tag{19b}$$

and the average number of input ports $\langle P_1 \rangle$ occupied on a cross connect at a node is

$$\langle P_1 \rangle = \left(\frac{D_1}{N} \right) [1 + \langle h \rangle] = \langle d \rangle [1 + \langle h \rangle]. \tag{19c}$$

Equations (19a)–(19c) are valid independent of the demand model, while as before the value of $\langle h \rangle$ is implicitly dependent upon the demand model. For the case of a mesh network with uniform demands, we substitute for $\langle d \rangle$ using (11c) to obtain

$$\langle P_1 \rangle = (N - 1) [1 + \langle h \rangle], \tag{19d}$$

where $\langle h \rangle$ may be approximated using (14b) or (14c).

For completeness, here we also compute the average number of two-way ports for a cross connect of the same network. The number of two-way terminations for a two-way demand is 2, one at each terminus. The average number of two-way through ports occupied is $2[1 + \langle h \rangle]$, and the total number of two-way ports occupied is

$$\begin{aligned} P_t &= \sum_{i<j}^{D_2} 2[1 + h_{ij}] = \frac{D_2}{D_2} \sum_{i<j}^{D_2} 2[1 + h_{ij}] \\ &= 2D_2 \langle 1 + h_{ij} \rangle = 2D_2 [1 + \langle h \rangle]. \end{aligned} \tag{19e}$$

Thus, the average number of two-way ports is

$$\langle P_2 \rangle = 2 (D_2/N) [1 + \langle h \rangle]. \tag{19f}$$

We observe that by substituting for D_2 using (10c)

$$\langle P \rangle \equiv \langle P_2 \rangle = \langle P_1 \rangle, \tag{20a}$$

which may be appreciated by again considering Fig. 3. This result is independent of the demand model and may also be structured to explicitly indicate the add, drop, and through ports. Considering Fig. 6, we write

$$\langle P\rangle \equiv \langle P_{\mathrm{ADD}}\rangle + \langle P_{\mathrm{DROP}}\rangle + \langle P_{\mathrm{THRU}}\rangle \tag{20b}$$

where

$$\langle P_{\mathrm{ADD}}\rangle = \langle P_{\mathrm{DROP}}\rangle = \langle d\rangle \tag{20c}$$

and

$$\langle P_{\mathrm{THRU}}\rangle = \langle d\rangle \left(\langle h\rangle - 1\right). \tag{20d}$$

Note that

$$\langle P_{\mathrm{ADD}}\rangle + \langle P_{\mathrm{DROP}}\rangle = 2\langle d\rangle \tag{20e}$$

which reminds us that every demand occupies both a termination-side port and a line-side port on each of the two cross connects at the opposite ends of the demand. Another common partitioning of ports is between termination-side ports and line-side ports. In this case, we write

$$\langle P\rangle \equiv \langle P_{\mathrm{TERM}}\rangle + \langle P_{\mathrm{LINE}}\rangle \tag{20f}$$

where

$$\langle P_{\mathrm{TERM}}\rangle = \langle P_{\mathrm{ADD}}\rangle = \langle d\rangle \tag{20g}$$

and

$$\langle P_{\mathrm{LINE}}\rangle = \langle P_{\mathrm{DROP}}\rangle + \langle P_{\mathrm{THRU}}\rangle = \langle d\rangle\langle h\rangle. \tag{20h}$$

In this analysis, for the average number of ports, we have not introduced the extra transmission capacity and extra cross-connect ports that are required for network survivability. As discussed previously, for single-link failure scenarios, the link, or line-side, capacity is increased by the fraction $\langle\kappa\rangle$. Thus, the total number of cross-connect ports for shared line-side restoration of mesh networks is obtained by introducing the extra capacity factor into (20h) and (19c), which yields

$$\langle P^{\kappa}\rangle = \langle d\rangle \left[1 + \left(1 + \langle\kappa\rangle\right)\langle h\rangle\right]. \tag{21a}$$

We note the same result is also obtained considering that the total number of ports is the sum of the number of channels carried on each of the links connected to the node and the number of channels terminating at the node. The former is given by the product of W^{o} and δ, and therefore

$$\langle P^{\kappa}\rangle = \langle d\rangle + \langle W^{o}\rangle \left(1 + \langle\kappa\rangle\right)\langle\delta\rangle. \tag{21b}$$

Using (13b) and (15b) and the definition of $\langle\kappa\rangle$, it can be shown that (21b) equates to (21a).

To appreciate how $\langle P\rangle$ scales with the number of nodes, we may consider (21) for uniform traffic in the limit when N is large compared with 1. In that limit and using (11c), (14c), and (17c) for $\langle d\rangle$, $\langle h\rangle$ and $\langle\kappa\rangle$, respectively, (21a) becomes

$$\langle P^{\kappa}\rangle \approx \left[\frac{\left(1 + \frac{2}{\langle\delta\rangle}\right)}{\sqrt{\langle\delta\rangle}}\right] N^{3/2}. \tag{22a}$$

For networks with $\langle\delta\rangle$ in the range of $3 \leq \langle\delta\rangle \leq 4$, the term in (21b) dependent upon $\langle\delta\rangle$ is within 15% of unity, and for $\langle\delta\rangle = 3.5$, the coefficient differs from 1 by less than 5%. Consequently, we observe that

$$\langle P^{\kappa}\rangle \approx N^{3/2}. \tag{22b}$$

Thus, if the number of nodes in the network is approximately 25, then the average number of ports required is about 125. When N is about 100, then $\langle P\rangle \sim 1000$; and when N is about 200, then $\langle P\rangle \sim 3000$. Similarly, the average traffic cross section carried on the route between adjacent nodes is

$$\langle W^{\kappa}\rangle \approx \frac{N^{3/2}}{\langle\delta\rangle} \tag{23}$$

when N is large compared with unity.

The average traffic handled by a cross-connect $\langle\chi\rangle$, measured in bits per second for example, is now computed straightforwardly from the average number of ports $\langle P\rangle$ and the communication bandwidth, either τ or B, associated with the basic unit of demand. Of course, the former corresponds to the case when the channel utilization is 100%, and the latter may correspond to a particular system increment or industry standard. Thus, we have

$$\langle\chi(\tau)\rangle \equiv \langle P\rangle\tau \tag{24a}$$

or

$$\langle\chi(B)\rangle \equiv \langle P\rangle B. \tag{24b}$$

These direct proportionalities are independent of the demand model. In Fig. 7, we graph the mean cross-connect traffic including idle restoration capacity for uniform demand as a function of the number of nodes N and total network traffic T.

For the larger example network of $N = 100$ nodes, $L = 171$ links, and $T = 5$ Tb/s traffic, the mean number of ports on a cross-connect including ports for restoration is estimated to be $\langle P^{\kappa}\rangle \cong 1061$. The corresponding mean cross-connect traffic is 1072 Gb/s. For the smaller example network of $N = 25$ nodes, $L = 42$ links, and $T = 1$ Tb/s traffic, the mean number of ports on a cross connect including ports for restoration is estimated to be $\langle P^{\kappa}\rangle \cong 141$. The corresponding mean cross-connect traffic is 469 Gb/s.

2) Variance and Standard Deviation: To compute the variance of P we must determine the number of ports required for the individual nodes. In the preceding sections, we have formulated expressions for the number of channels on the individual links, namely (15d)–(15g), (16b), and (17d). Consequently, it is necessary only to add the termination-side channels to the sum of the channels on the δ_i links connected to an individual node i to obtain the sum of the ports, ${P^{\kappa}}_i$. Formally, we may write

$$P_i^{\kappa} = d_i + \sum_{j}^{\delta_i} W_{ij}^{\kappa}. \tag{25a}$$

Hence, the variance of P^{κ} may be computed using this expression and the definition of the variance, (13d). In the spirit of clarifying the dependencies of the variance of P^{κ}, we now illustrate an example where the local extra capacity for restoration

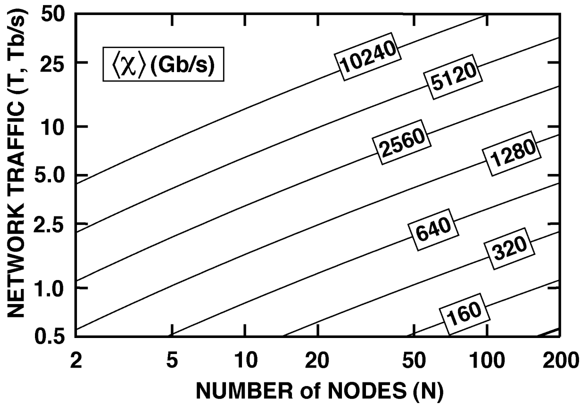


Fig. 7. Mean cross-connect traffic. The mean traffic entering a cross-connect $\langle\chi^\kappa(N,T)\rangle$ for uniform demand with restoration is graphed as a function of the number of nodes N and total two-way traffic T under the constraint $\langle\delta\rangle = 3.5$ using a contour plot.

is specified by (17d). In this scenario, the number of ports on a local cross connect is

$$P^\kappa{}_i \cong 2d_i + \left[\frac{d_i}{\langle\delta\rangle} + W_{B/T} + \frac{\langle W^o\rangle}{\langle\delta\rangle}\right]\delta_i + \langle W^o\rangle \tag{26a}$$

where for the total extra capacity associated with ports at node i, we have used the approximation

$$\kappa_i = \sum_j^{\delta_i} \kappa_{ij} \cong 1 + \frac{\delta_i}{\langle\delta\rangle}. \tag{26b}$$

Considering (26a), we observe there is a correlation between $P^\kappa{}_i$ and δ_i that is moderated by the variations in W_T. The variance of P^κ for uniform demand is given by

$$\sigma^2(P^\kappa) \cong \left[\frac{\langle d\rangle}{\langle\delta\rangle} + \langle W_{B/T}\rangle + \frac{\langle W^o\rangle}{\langle\delta\rangle}\right]^2 \sigma^2(\delta) + \langle\delta\rangle^2\sigma^2(W_T). \tag{26c}$$

If instead we use (17g) to specify the extra capacity on a link, then the total extra capacity associated with ports at node i is

$$\kappa_i \cong \langle\delta\rangle\left[\frac{1}{\delta_i} + \left\langle\frac{1}{\delta}\right\rangle\right] \tag{27a}$$

and the total number of ports P^κ is

$$P^\kappa{}_i \cong 2d_i + \left[\frac{d_i}{\langle\delta\rangle} + W_{B/T}\right]\delta_i + \frac{\langle W^o\rangle\langle\delta\rangle}{\delta_i} + \langle W^o\rangle\langle\delta\rangle\left\langle\frac{1}{\delta}\right\rangle. \tag{27b}$$

In this case, there is a contribution to the number of ports from the extra capacity $(1/\delta_i)$ that is anti-correlated with the main term that is proportional to δ_i. Thus, we expect the variance of P^κ in this scenario for the extra capacity (17g) to be somewhat less than the variance obtained using the first form (17d).

To illustrate the variance of P^κ, we consider the situation in which the variance of W_T is small and may be neglected. In this case, (26c) reduces to

$$\sigma(P^\kappa) = \langle W^o\rangle\left[1 + \frac{1}{\langle\delta\rangle} - \frac{1}{\langle h\rangle}\right]\sigma(\delta). \tag{28a}$$

For our example network of $N = 100$ nodes and $L = 171$ links, the mean and standard deviation of the degree of nodes is $\langle\delta\rangle = 3.4$ and $\sigma(\delta) = 1.1$. Using (28a), the standard deviation of the number of ports on a cross connect attributable to the variance of the degrees of nodes is estimated to be $\sigma(P^\kappa) \cong 219$. Recall the mean number of ports including restoration capacity was estimated to be $\langle P^\kappa\rangle \cong 1061$ for this network. We expect that the fractional deviations for our smaller example network of $N = 25$ nodes and $L = 42$ links will be similar, as the statistics of the degrees of nodes are nearly the same by design. Again using (28a), the standard deviation of the number of ports on a cross-connect for this smaller network is estimated to be $\sigma(P^\kappa) \cong 23$. For this smaller network, the mean number of ports including restoration capacity was estimated to be $\langle P^\kappa\rangle \cong 141$.

In summary, in this and the preceding section, we have shown that the network global expectation model can be used to understand and predict the mean and variability of the number channels carried on links and present at the nodes, including the effects resulting from network survivability. The reader will appreciate that while we have applied the model to the case of uniform demand in this section on the variance of the number of ports, the methodology is directly applicable to other demand profiles.

J. Percentage Add/Drop

1) Mean Value: Another important characteristic of the network is the percentage of add and drop traffic at a node. Referring to Fig. 6 and the one-way input ports on the cross connect, we observe that the average number of input ports occupied by traffic being either added or dropped at the node is

$$\langle P_{\mathrm{ADD}}\rangle + \langle P_{\mathrm{DROP}}\rangle = \frac{D_1}{N} + \frac{D_1}{N} = \frac{2D_1}{N}. \tag{29a}$$

The average number input ports occupied by traffic passing through the node is

$$\langle P_{\mathrm{THRU}}\rangle = \frac{D_1(\langle h\rangle - 1)}{N}. \tag{29b}$$

By definition, the average ratio of the number of local A/D ports to local total ports is

$$\langle\rho\rangle \equiv \frac{1}{N}\sum_n^N \frac{(P_{\mathrm{ADD}} + P_{\mathrm{DROP}})_n}{P_n} \tag{30a}$$

which may be computed by substituting expressions for both numerator and denominator. However, another practical and useful definition of the A/D ratio average is the ratio of the network total number of A/D ports to network total ports. In this second case

$$\begin{aligned}\langle\rho'\rangle &= \frac{N(\langle P_{\mathrm{ADD}}\rangle + \langle P_{\mathrm{DROP}}\rangle)}{N(\langle P_{\mathrm{ADD}}\rangle + \langle P_{\mathrm{DROP}}\rangle + \langle P_{\mathrm{THRU}}\rangle)} \\ &= \frac{(\langle P_{\mathrm{ADD}}\rangle + \langle P_{\mathrm{DROP}}\rangle)}{\langle P\rangle}\end{aligned} \tag{30b}$$

and therefore

$$\langle\rho'\rangle = \frac{2}{[1 + \langle h\rangle]}. \tag{30c}$$

Note that we have derived this relationship between $\langle\rho'\rangle$ and $\langle h\rangle$ without reference to a model for the demands D_1; consequently, it is a general result and not restricted to the case of uniform demands. If we explicitly account for extra capacity for line-side restoration, then the ratio average $\langle \rho'_\kappa \rangle$ of the number of A/D ports to total ports (21) is

$$\langle \rho'^\kappa \rangle = \frac{2}{[1 + (1 + \langle \kappa \rangle) \langle h \rangle]}. \tag{30d}$$

The estimated A/D ratios for the example network of $N = 100$ nodes and $L = 171$ links without and with extra capacity for restoration are $\langle \rho' \rangle \cong 0.28$ and $\langle \rho'^\kappa \rangle \cong 0.19$ using (30c) and (30d), respectively. In comparison, the estimated A/D ratios for the example network of $N = 25$ nodes and $L = 42$ links without and with extra capacity for restoration are $\langle \rho' \rangle \cong 0.49$ and $\langle \rho'^\kappa \rangle \cong 0.34$ using (30c) and (30d), respectively. This trend of the fraction of the through traffic increasing as the number of nodes is increased is a general characteristic of networks having a mean degree of node less than $N - 1$ and fully interconnected terminal-to-terminal demand. In the limit, when N is large compared with 1 and the average degree of node is in the range $3 \le \langle \delta \rangle \le 4$, the total number of ports is given by (22b) and the A/D ratio average becomes

$$\langle \rho'^\kappa \rangle \approx \frac{2}{\sqrt{N}}. \tag{30e}$$

Thus, for a mesh network of 25 nodes with shared restoration capacity, the ratio of A/D to through channels is approximately 40% on average, and the percentage decreases as the number of nodes increases. Of course, this estimate is for the average node, and the percentage for a particular node can be larger or smaller, depending upon the details of the network demand and topology.

On a separate note related to the A/D ratio, it is also worth pointing out that (30c) may be inverted to express $\langle h \rangle$ as a function of $\langle \rho' \rangle$, viz.

$$\langle h \rangle = \left[\frac{2}{\langle \rho' \rangle} - 1 \right]. \tag{31}$$

Like (30c), (31) is a general result and is not a function of the demand model.

2) Variance and Standard Deviation: The ratio of the A/D traffic to total traffic for an individual node may be formulated using (25) and (29a). For example, if we consider the case when $\sigma(W_T)$ is negligible, the result using (17d) for the extra capacity is

$$\rho_i^\kappa = \frac{2\langle d \rangle}{P_i^\kappa} = \frac{2}{1 + \left(1 + \frac{2}{\langle \delta \rangle}\right) \sqrt{\frac{\langle d \rangle}{\delta} \frac{\delta_i}{\langle \delta \rangle}}}. \tag{32a}$$

When N is large compared with 1 and $\langle \delta \rangle$ is in the range of $3 \le \langle \delta \rangle \le 4$, we may approximate (32a) by

$$\rho^\kappa{}_i = \left(\frac{2}{\sqrt{N}} \right) \left[\frac{\langle \delta \rangle^{\frac{3}{2}}}{\left(1 + \frac{2}{\langle \delta \rangle}\right)} \right] \left(\frac{1}{\delta_i} \right), \tag{32b}$$

and so in this case

$$\frac{\sigma(\rho^\kappa)}{\langle \rho^\kappa \rangle} \approx \frac{\sigma\left(\frac{1}{\delta}\right)}{\left\langle \frac{1}{\delta} \right\rangle}. \tag{32c}$$

In addition,

$$\frac{\rho^\kappa_{\min/\max}}{\langle \rho^\kappa \rangle} \approx \frac{\langle \delta \rangle}{\delta_{\max/\min}}. \tag{32d}$$

Thus, given that δ_i may range from 2 to 8, we conclude that the A/D ratio can conceivably range from one half to two times the mean value.

IV. Network Cost

A. Node and Link Architecture

In the previous section, we have derived expectation values for the quantities of key NEs and NE subsystems required to carryout a basic cost analysis for a transport network. In this section, we will introduce the concept of the cost structure of NEs in relation to both the NEs and NE subsystems. With an assumed cost structure, we may then compute the total cost of the network as well as categories of costs, such as for transmission and bandwidth management. We also illustrate by example how we may compare the network costs using different combinations of technology, such as electronic and optical bandwidth management, using the network global expectation model.

For the purpose of outlining the general principles of computing network costs using the network global expectation model, here we consider rudimentary cost structures for the optical line system (OLS), EXC, and OXC. The architecture of these systems from a perspective near a node is illustrated in Fig. 8. Client-side traffic enters the network at a node via the EXC where it is groomed, i.e., switched and multiplexed, into the fundamental units of internodal bandwidth destined for specific nodes of the network. The groomed output channels from the EXC then enter the OXC, where they are directed to line systems placed along the internodal links of the network according to the traffic routing scheme determined by either a centralized or distributed management system. In the architecture considered here, the interfaces between NEs are optical translators (OTs), which ensure that the cost comparisons are under conditions of fixed network capability (features) and network performance.

B. Transmission Cost Structure

A cost structure often used for optical fiber transmission is the average cost of transporting bandwidth (B) over distance (s). Here, we represent this cost structure as a cost coefficient, which we denote as γ_{B-s}. The units of γ_{B-s} are dollars per gigabit per second per kilometer ($/Gb/s/km). According to Gawrys, an approximate value for network transmission cost of a two-way channel is

$$\gamma_{B-s} \approx 30\ \$/\text{Gb/s/km} \tag{33}$$

based on historical data and projections [18].

Considering this cost structure, the individual and mean cost of a transmission link of a survivable mesh network are

$$c_i = \gamma_{B-s}\, \beta\, s_i \tag{34a}$$

$$\langle c_l \rangle = \gamma_{B-s} \langle \beta\, s \rangle \cong \gamma_{B-s} \langle \beta \rangle \langle s \rangle \tag{34b}$$

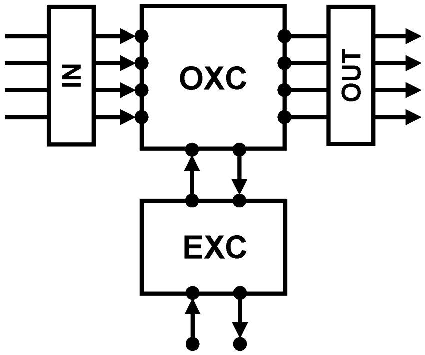


Fig. 8. Bandwidth management architecture using both OXCs and EXCs. For a cross connect using only electronic bandwidth management, a single EXC replaces the combination of OXC and EXC.

where for the model of uniform demand under present consideration $\langle\beta\rangle$ is given by (16) with $\langle\kappa\rangle$ given by (17c) and $\langle s\rangle$ is the expectation value of the link length. The expectation value of the link length $\langle s\rangle$ is given by

$$\langle s\rangle \equiv \frac{1}{L}\sum_{l}^{L} s_l \tag{35a}$$

where the set $\{s\}$ are the physical lengths of the individual links. If the link lengths are known, then the expectation value $\langle s\rangle$ is quickly computed. Here, for purposes of illustration, without introducing a specific set of link lengths, we note that for two-dimensional mesh networks, to a good approximation the average link length scales inversely with the square root of the number of nodes and is proportional to the square root of the geographic area A covered by the network. Thus

$$\langle s\rangle \cong \frac{\sqrt{A}}{(\sqrt{N}-1)}. \tag{35b}$$

The total cost of transmission is

$$C_{\mathrm{TRANS}} = L\langle c_l\rangle \tag{36a}$$

where it should be clear that C_{TRANS} is an analytic function of only the independent input network variables (N, the number of nodes; L, the number of links; T, the total ingress/egress traffic; and A, the geographic area covered by the network) and therefore is easily computed. Consequently, when N is large compared with 1, and $\langle\delta\rangle$ is in the range of $3 \le \langle\delta\rangle \le 4$, C_{TRANS} may be approximated by

$$C_{\mathrm{TRANS}} \approx \gamma_{B-s} T\sqrt{A}. \tag{36b}$$

At present, the yearly time-averaged traffic carried by a combined voice and data backbone network in the continental United States is approximately 1 Tb/s [19], [20]. The daily and annual peak traffic load that the network must support is estimated to be ~5 × the average traffic [20]. Thus, as an example, we consider $T = 5$ Tb/s. The geographic area of the continental United States is approximately $A = 8 \times 10^6$ km^2. Thus, the approximate cost of transmission system equipment C_{TRANS} to support the present traffic is approximately 400 $M.

The approximate cost of transmission represented by (36b) is obviously an over simplification as it contains no dependency on the number of links. That behavior is not because of a shortcoming of the global network expectation model but, rather, is attributed to our assumption of the cost structure, (33) and (34). Clearly a more realistic model of the cost structure for the link should include an explicit dependency upon the cost of trenching, the cost of optical fiber cable, the cost of end terminals, the cost of OTs, the cost of amplifiers, and the cost of amplifier pumps, for example. Realizing this, a refined cost structure for a link takes the form

$$c_i = \gamma_{t0} + \gamma_{t1}\,\tau\,W_i + \gamma_{t2}\,s_i + \gamma_{t3}\,\tau\,W_i s_i. \tag{37a}$$

The expectation value for the cost of a link is then

$$\langle c_1\rangle = \frac{1}{L}\sum_{i}^{L} c_i = \frac{1}{L}\sum_{i}^{L}\{\gamma_{t0}+\gamma_{t1}\,\tau W_i+\gamma_{t2}\,s_i+\gamma_{t3}\,\tau W_i s_i\} \tag{37b}$$

where the first term containing γ_{t0} reflects fixed costs for a link, such as the cost of the terminal equipment bays; the second term containing γ_{t1} includes costs that depend directly upon the number of channels carried, such as the number of OTs, the third term containing γ_{t2} includes costs that depend upon the distance traversed, such as the cost of trenching, cost of fiber, and the cost of amplifiers; and the fourth term containing γ_{t3} includes contributions that grow as the product of distance and wavelength, such as the cost of growth pumps and premium for specialized high capacity, long-distance fiber, e.g., dispersion-managed cable. The total cost of transmission is then

$$C_{\mathrm{TRANS}} = L\langle c_l\rangle = L\left\{\gamma_{t0}+\gamma_{t1}\,\tau\langle W\rangle+\gamma_{t2}\,\langle s\rangle+\gamma_{t3}\,\tau\langle Ws\rangle\right\}. \tag{37c}$$

Of the expectation values contained in (37), in this paper, we have previously computed all except for $\langle Ws\rangle$. As we have observed before, the number of channels on a link for the case of uniform demands is nearly independent of the particular link. Thus, to a good approximation $\langle Ws\rangle = \langle W\rangle\langle s\rangle$ and the total cost of transmission is

$$C_{\mathrm{TRANS}} \cong L\left\{\gamma_{t0} + \gamma_{t1}\,\tau\langle W\rangle + \gamma_{t2}\,\langle s\rangle + \gamma_{t3}\,\tau\langle W\rangle\,\langle s\rangle\right\}. \tag{37d}$$

The above approximation is further validated when we consider that under real-world circumstances, the coefficient γ_{t3} is small compared with the other coefficients, and rarely are the optical line systems loaded to their maximum channel carrying capacity. In this case, to gain a better appreciation for how the total transmission cost depends upon the basic network variables, we consider dropping the last term. Upon substituting for the remaining expectation values in (37d), the cost of transmission is then

$$C_{\mathrm{TRANS}}(N,T) \cong \frac{1}{2}\left[\gamma_{t0} + \frac{\gamma_{t2}\langle\delta\rangle N\sqrt{A}}{(\sqrt{N-1})}\right. \\ \left. + \gamma_{tl}\left[\frac{\sqrt{A}\left(1+\frac{2}{\langle\delta\rangle}\right)}{\sqrt{\langle\delta\rangle-1}}\right]T. \right] \tag{37e}$$

Here, the fixed startup costs, i.e., those independent of the traffic carried T, are evident in the first term, which is proportional to N or L ($L = N\langle\delta\rangle/2$, (13b)). We leave it to the reader to apply (37d) to specific network designs by substituting values for the cost structure coefficients γ.

C. Bandwidth Management Architectures and Cost Structure

1) Electronic Bandwidth Management Only: The network global expectation model provides the flexibility and ease of implementation to compute the NE variables and total network costs for a wide range of network sizes, total traffic, and a variety of architectural options. Here, we illustrate how the costs for two different models of bandwidth management at the network nodes may be constructed. We first consider the case when an EXC is used for both sub-rate grooming and cross-connect functions. In this case, the total cost of bandwidth management is the cost of the EXC, and so

$$C_{\mathrm{BWM}} = C_{\mathrm{EXC}}. \tag{38}$$

The total cost of the EXCs may be written in terms of the expectation value of the cost of the nodes as

$$C_{\mathrm{EXC}} = \langle c_{\mathrm{EXC}}\rangle N, \tag{39a}$$

which follows directly from (8). An estimate of the current cost of high-speed electronic switching engines is

$$\gamma_{\mathrm{ep}} \approx 1\ \text{\$K/Gb/s} \tag{39b}$$

which corresponds to a cost structure of the local EXC of

$$c_{\mathrm{EXC}} = \gamma_{\mathrm{ep}}\chi(\tau). \tag{39c}$$

The corresponding expectation value is

$$\langle c_{\mathrm{EXC}}\rangle = \gamma_{\mathrm{ep}}\,\langle\chi(\tau)\rangle = \gamma_{\mathrm{ep}}\,\tau\,\langle P\rangle \tag{39d}$$

having made use of (24a). Substituting for $\langle c_{\mathrm{EXC}}\rangle$ in (39a) and using (12a) and (21a), we have

$$C_{\mathrm{EXC}} = \langle c_{\mathrm{EXC}}\rangle N = 2\gamma_{\mathrm{ep}}T\left[(2+\langle\kappa\rangle)\,\langle h\rangle\right]. \tag{39e}$$

Note that we may also construct a more refined form for the cost structure of the EXC, or IPRs, that includes a start-up term and a growth term, viz.

$$c_{\mathrm{EXC}} = \gamma_{e0} + \gamma_{e1}\,\chi_\tau. \tag{39f}$$

In this case

$$C_{\mathrm{EXC}}(N,T) = \langle c_{\mathrm{EXC}}\rangle N = \gamma_{e0}N + 2\left[(2+\langle\kappa\rangle)\,\langle h\rangle\right]\gamma_{e1}T. \tag{39g}$$

We note these expressions for costs are valid independent of the demand model.

2) Electronic and Optical Bandwidth Management: Here, we consider a single-tier network using both optical and electronic bandwidth management. By this, we mean that all traffic passes through the optical layer cross connect and additionally all terminating traffic also passes through an electronic layer fabric for the purpose of channel grooming. Such an architecture is attractive when the cost of an optical port is significantly less than the cost of an electronic port for a given data rate. The total cost for bandwidth management is thus

$$C_{\mathrm{BWM}} = C_{\mathrm{EXC}} + C_{\mathrm{OXC}}. \tag{40}$$

In the following subsection, we construct the individual terms for the EXC and OXC costs.

a) Cost of electronic ports for client-side traffic: As before, we assume that the cost of the electronic switch consists of a start-up term and a term proportional to the traffic handled; however, here only the terminating traffic traverses the EXC. Thus, the mean cost of an EXC is

$$\langle c_{\mathrm{EXC}}\rangle = \gamma_{e0} + \gamma_{e1}\,\tau\langle P_{\mathrm{ADD}} + P_{\mathrm{DROP}}\rangle = \gamma_{e0} + \gamma_{e1}2\,\tau\langle P_{\mathrm{ADD}}\rangle \tag{41a}$$

which may be rewritten as

$$\langle c_{\mathrm{EXC}}\rangle = \gamma_{e0} + 4\gamma_{e1}\,T/N \tag{41b}$$

using (12) for τ. Consequently

$$C_{\mathrm{EXC}} = \gamma_{e0}N + 4\gamma_{e1}\,T. \tag{41c}$$

b) Cost of optical ports for through and A/D traffic: The total cost of OXCs using the network global expectation formalism is

$$C_{\mathrm{OXC}} = \langle c_{\mathrm{OXC}}\rangle N. \tag{42a}$$

An estimate of the current cost of high-speed optical switching engines is

$$\gamma_{\mathrm{op}} \approx 2.5\ \text{\$K/port}. \tag{42b}$$

Based on this cost structure and the architecture under consideration, which specifies that both through and client-side traffic pass through the OXCs, the individual and mean OXC costs may be expressed as

$$c_{\mathrm{OXC}} = \gamma_{\mathrm{op}}\,P \tag{42c}$$

and so

$$\langle c_{\mathrm{OXC}}\rangle = \gamma_{\mathrm{op}}\langle P\rangle. \tag{42d}$$

Substituting variables to obtain an expression that is independent of the demand model, the total cost of the OXCs is

$$C_{\mathrm{OXC}}(N) = \langle c_{\mathrm{OXC}}\rangle N = 2\gamma_{\mathrm{op}}\,D(N)\left[(2+\langle\kappa\rangle)\,\langle h\rangle\right] \tag{42e}$$

where $D(N)$ is the number of two-way demands.

As in the other examples, we may also consider a cost structure for the OXC consisting of a start-up term and a growth term, such as

$$c_{\mathrm{OXC}} = \gamma_{o0} + \gamma_{o1}\,P. \tag{42f}$$

In this case, the mean and total cost of the OXCs are

$$\langle c_{\mathrm{OXC}}\rangle = \gamma_{o0} + \gamma_{o1}\langle P\rangle \tag{42g}$$

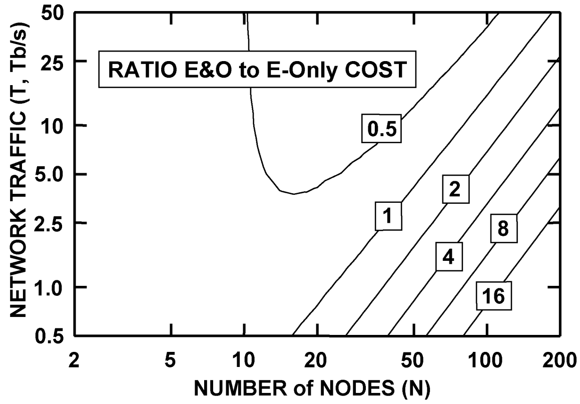


Fig. 9. Illustrative comparison of bandwidth management costs. The total cost of bandwidth management using the combination of optical and EXCs is compared with the total cost of bandwidth management using only an EXC by plotting their ratio as a function of number of nodes N and two-way traffic T. In the case of the E&O architecture, it is assumed that all traffic follows through the optical switch fabric and additionally that all terminating traffic flows through the electronic switch fabric. These calculations are for uniform demand with restoration under the constraint $\langle\delta\rangle = 3.5$. The cost structures (γ) used for the OXCs and EXCs for this example are \$2.5 K/port and \$1 K/Gb/s, respectively. Note that these cost structures and values are rudimentary, intended to be illustrative, and should not be interpreted as definitive.

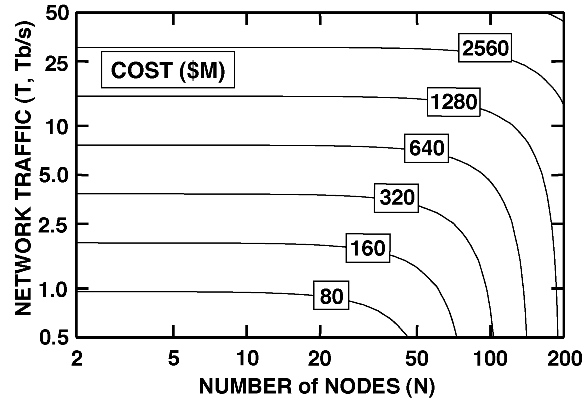


Fig. 10. Total network equipment cost. The sum of transmission and bandwidth management equipment costs $C_T(N,T)$ is graphed as a function of the number of nodes N and total two-way traffic T using a contour plot. Again, the calculations are for uniform demand with restoration under the constraint $\langle\delta\rangle = 3.5$. The cost structures used for the optical line systems, EXCs, and OXCs are \$30/Gb/s/km, \$1 K/Gb/s, and \$2.5 K/port, respectively. The reader is again cautioned that these cost structures and values are intended only to illustrate the capabilities and possibilities of the global expectation model.

and

$$C_{\mathrm{OXC}}(N) = \langle c_{\mathrm{OXC}}\rangle N = \gamma_{o0}N + 2\gamma_{o1}D(N)\left[(2+\langle\kappa\rangle)\langle h\rangle\right]. \tag{42h}$$

Summing the electronic and optical bandwidth management costs, we have

$$C_{\mathrm{BWM}}(N,T) = (\gamma_{e0}+\gamma_{o0})\,N + 4\gamma_{e1}T + 2\gamma_{o1}D(N) \times \left[(2+\langle\kappa\rangle)\langle h\rangle\right]. \tag{43}$$

3) Comparison of Costs for Example Node Architectures: As an illustration of the application of the network global expectation model, we compare the total costs for bandwidth management for the two-node architecture examples just described, namely electronic plus optical bandwidth management and electronic-only bandwidth management, as a function of the number of nodes N and traffic T for a fixed mean degree of node. The results of the calculations using the coarse cost structures for the EXC and OXC costs, (39b) and (42b), are graphed in Fig. 9. We observe that the model may be used to identify the region of the network parameter space where optical layer cross connects may be introduced in conjunction with electronic layer cross connects, or IPRs, to economic advantage. The model accounts not only for the different characters of the cost structures as a function of traffic, but also the changing ratio of A/D to through traffic as the number of nodes and links change.

We observed for a fixed value of the number of nodes for N greater than about 15 that the total cost of bandwidth management using the electrical and optical architecture decreases and becomes less than the cost of the electronic-only solution as the total traffic increases. This is attributed to the assumption that the cost of an optical switch port is independent of channel bit rate, while the cost of an electronic switch port is directly proportional to the channel bit rate. We also observe for fixed total network traffic that the cost of the electronic and optical solution increases and becomes more expensive than the electronic-only solution as the number of nodes is increased and the mean degree of the nodes is held constant. This is because the traffic demand bit rate τ decreases as the number of nodes is increased for fixed mean degree of the nodes (see Fig. 4) and, consequently, below some channel bit rate, the fixed cost of an optical switch port becomes more expensive than an electronic switch port.

Of course, the details of the cost crossover depend upon the particulars of the technology price points (cost structure and coefficients), and consequently the particular graph of Fig. 9 is intended only to demonstrate the capabilities and possibilities of the global expectation model and not to make a definitive recommendation. The reader is also cautioned that here we have implicitly assumed via the cost structures that the respective cross connect technologies are capable of providing the required switch and backplane capacities. In the absence of more refined cost structures that account for these limitations, one may use other equations and graphs of the model, such the total number of required ports (21b) or the mean cross-connect traffic (Fig. 7), to identify regions of the network traffic-node space that are beyond the capabilities of a particular architecture or technology.

D. Total Network Costs

1) Static Network Cost: The total network cost can be computed by summing the cost for transmission and bandwidth management using the formulas we have derived. For completeness, we explicitly restate (4) here as

$$C_T = C_{\mathrm{TRANS}} + C_{\mathrm{BWM}}. \tag{44}$$

Clearly, a useful attribute of the model is that the relative cost of transmission and bandwidth management can easily and quickly be determined.

To illustrate the utility of the network global expectation model, in Fig. 10, we present a calculation of the total cost of a

TABLE I
KEY ANALYTIC EXPRESSIONS. KEY RESULTS OF PRESENT FORMULATION OF THE NETWORK GLOBAL EXPECTATION MODEL AND THEIR CORRESPONDING TYPES AND DOMAINS OF APPLICABILITY ARE SUMMARIZED

Variable	Relationship	Equation	Type	Applicability
Demand Bit-Rate	$\tau \equiv T/D$	(12a)	exact	independent of demand model
Degree of Node	$\langle\delta\rangle = 2L/N$	(13b)	exact	independent of demand model
Inter-nodal Hops	$\langle h\rangle \cong [(N - 2)/(\langle\delta\rangle - 1)]^{1/2}$	(14c)	semi-empirical	minimum hop routing, location-independent demand, nominally planar topology
Channels on Link	$\langle W^o\rangle = \langle d\rangle \langle h\rangle / \langle\delta\rangle$	(15b)	exact	independent of demand model without survivability
Channels on Link	$\langle W^\kappa\rangle \equiv \langle W^o\rangle (1 + \langle\kappa\rangle)$	(16a)	exact	independent of demand model with survivability
Restoration Capacity	$\langle\kappa\rangle \cong 2/\langle\delta\rangle$	(17c)	semi-empirical	location-independent demand, restricted working/restoration differential path length
Link Traffic	$\langle\beta\rangle \equiv \langle W\rangle\tau$	(18a)	exact	independent of demand model
Channels at XC	$\langle P^\kappa\rangle = \langle d\rangle + \langle W^o\rangle (1 + \langle\kappa\rangle) \langle\delta\rangle$	(21b)	exact	independent of demand model
Cross-Connect Traffic	$\langle\chi\rangle \equiv \langle P\rangle \tau$	(24a)	exact	independent of demand model
Add/Drop Ratio	$\langle\rho'\rangle = 2/[1 + \langle h\rangle]$	(30c)	exact	independent of demand model
Link Length	$\langle s\rangle \cong \sqrt{A}/(\sqrt{N} - 1)$	(35b)	semi-empirical	nominally planar topology, minimum hop routing, location-independent demand

mesh network with uniform demand as a function of the number of nodes N and total traffic T. The results are for the case where the nodal bandwidth manager consists of a combination of OXCs and EXCs and the geographic area corresponds to the continental United States. In the accounting we have used (33), (34), (39b), (39c), (42b), (42c) for the cost structure of the transmission links, EXCs, and OXCs, respectively.

Among the features that may be observed by considering Fig. 10 is the impact of the cost of bandwidth management as the number of nodes increases. A qualitatively similar result is obtained for the case of electronic-only bandwidth management. Considering (22b) for total number of cross-connect ports and (30e) for the A/D ratio, we interpret the large cost for large N to be a consequence of the single-tier architecture. In effect, single-tier (flat) networks can not practically scale to a very large number of nodes because as the number of nodes increases, an increasing fraction of the traffic processed at each node is through traffic destined for other nodes. It is for this reason that the voice and packet networks are organized hierarchically based on geographic communities. The underlying phenomenon may also be the driving factor behind the more broadly observed scaling behavior of networks and biological systems [21]. Clearly, there are performance and operational tradeoffs between single-tier and multi-tier networks, and network operators will adjust the number of nodes and architecture in the backbone depending upon the costs for transmission and bandwidth management; changing cross-connect, line-system, and technology price points; and the evolution of traffic demand.

2) Refinement of Cost Structure and Evolution of Network Cost: Here, we take the opportunity to mention additional ways in which the cost structure may be refined and the model may be applied. First, the cost structure may be modified to account for the real-world implementation limits affecting maximum system capacities. Examples of such constraints are the maximum number of channels or wavelengths an optical line system is engineered for, or the maximum throughput of a switch fabric or backplane in the case of a cross connect or router. Such hard bounds to NE capacity occur for any physical realization and have the effect of introducing quantum steps in the cost structure. When the required capacities exceed the system capabilities, additional systems are generally deployed in parallel, and additional corresponding start-up costs are incurred. Having developed a framework for the evaluation of the variances and distribution functions of key network variables previously in this paper, we have provided the foundation to estimate the number of additional systems that are required given the network requirements and system bounds. Note too that in some instances the result of introducing these additional systems is to effectively increase the number of links or nodes of the network. Second, the model may be used for sensitivity analyses of the dependency of requirements and costs upon primary and secondary network and NE variables. Third, the network global expectation model may be used to compute the constituent and total network costs as a function of time. This requires only a model for how the total network traffic, number of nodes and links, and technology costs are expected to change, such as have been described in other works [18], [20].

V. Conclusion and Summary

Here, we have described a network global expectation model as a comprehensive and structured framework for estimating the number of NEs, NE characteristics, and costs of communication networks using analytic formulas. The model includes the calculation of both the mean value and variance of all key network quantities and may be applied to a wide range of topologies, architectures, and demand profiles. Currently, we have formulated the general approach, have applied it to single-tier mesh networks and location-independent demands, and have also shown that many of the results are valid and applicable independent of the demand model. For uniform demands, we have shown that the number of nodes, the degrees of the network nodes, the total ingress/egress traffic, the geographic extent of the network, and the equipment cost structures are sufficient to estimate the network variables and costs of interest. We have also formulated either exact or semi-empirical functions and closed-form expressions for the network variables, which are easily incorporated into software spreadsheet calculators. For the convenience of the reader, key results of the model are summarized in Table I.

This analytic tool naturally and accurately relates the global (network) and local (NE) views of the communication system and thereby can quickly provide insight and roughly correct results for preliminary network evaluation and design. Further, it can provide valuable guidance in the areas of NE feature requirements, costs, sensitivity analyses, scaling performance, comparisons, product definition and application domains, and product and technology roadmapping. It is adaptable to both increasing and decreasing levels of detail and sophistication of the cost structures. Because of the analytic nature of the model, the estimates of quantities may be computed much faster than is possible with detailed routing solvers, and so the model is ideally suited to network analyses in dynamic operating and technological environments. We suggest that the uncomplicated and transparent accounting of NEs, systems, and costs inherent in the model can constitute a framework for the cooperative exchange of critical planning information on evolving network needs across the many sectors of the communication business. In future work, we plan to refine and extend this approach to a wider set of networks, architectures, demand profiles, and cost structures.

Acknowledgment

The author thanks V. Arya, G. Austin, D. Einstein, T. Mueller, and I. Saniee for their critical and constructive feedback on the value, applicability, and features of the network global expectation model during the early phase of this work. In addition, the author is grateful to S. Baroni, D. Levy, and G. Wilfong for their careful reading of this manuscript, thoughtful questions, and valuable comments and suggestions. Finally, the author acknowledges R. Alferness, G. Austin, and D. Bishop for their encouragement and support to pursue the formulation of this network analysis model and tool.

References

[1] *Optical Fiber Telecommunications IVA & IVB*, I. P. Kaminow and T. Li, Eds., Academic, San Diego, CA, 2002.

[2] D. Al-Salameh, S. K. Korotky, D. S. Levy, T. O. Murphy, S. H. Patel, G. W. Richards, and E. S. Tentarelli, "Optical switching in transport networks: Applications, requirements, architectures, technologies, and solutions," in *Optical Fiber Telecommunications IVA*, I. P. Kaminow and T. Li, Eds. San Diego, CA: Academic, 2002, pp. 295–373.

[3] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: An approach to high bandwidth optical WAN's," *IEEE Trans. Commun.*, vol. 40, pp. 1171–1182, July 1992.

[4] B. Mukherjee, S. Ramamurthy, D. Banerjee, and A. Mukherjee, "Some principles for designing a wide-area optical network," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, 1994, pp. 110–119.

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

[6] S. Baroni and P. Bayvel, "Wavelength requirements in arbitrarily connected wavelength-routed optical networks," *J. Lightwave Technol.*, vol. 15, pp. 242–251, Feb. 1997.

[7] S. Baroni, S. K. Korotky, and P. Bayvel, "Wavelength interchange in multi-wavelength optical transport networks," in *Tech. Dig. 23rd Eur. Conf. Opt. Commun.*, vol. 3, Edinburgh, 1997, pp. 164–167.

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

[9] B. T. Doshi, S. Dravida, P. Harshavardhana, O. Hauser, and Y. Wang, "Optical network design and restoration," *Bell Labs. Tech. J.*, vol. 4, no. 1, pp. 58–84, 1999.

[10] S. Baroni, P. Bayvel, R. J. Gibbens, and S. K. Korotky, "Analysis and design of resilient multi-fiber wavelength-routed optical transport networks," *J. Lightwave Technol.*, vol. 17, pp. 743–758, May 1999.

[11] R. D. Doverspike, S. J. Phillips, and J. R. Westbrook, "Transport network architectures in an IP world," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, 2000, pp. 305–314.

[12] R. D. Davis, K. Kumaran, G. Liu, and I. Saniee, "SPIDER: A simple and flexible tool for design and provisioning of protected lightpaths in optical networks," *Bell Labs Tech. J.*, vol. 6, no. 1, pp. 82–97, 2001.

[13] A. W. Goodman and J. S. Ratti, *Finite Mathematics With Applications*. New York: Macmillan, 1971, pp. 367–399.

[14] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vettering, *Numerical Recipes, the Art of Scientific Computing*. New York: Cambridge Univ. Press, 1986, p. 456.

[15] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*. New York: McGraw-Hill, 1969.

[16] M. Dwass, *Probability and Statistics*. New York: W. A. Benjamin, 1970.

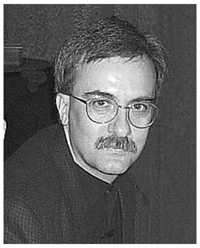
[17] J. Strand, "Optical network architecture evolution," in *Optical Fiber Telecommunications IVB*, I. P. Kaminow and T. Li, Eds. San Diego, CA: Academic, 2002, p. 111.

[18] G. W. Gawrys, "Delivering the promise of optical networking," presented at the Nat. Fiber Opt. Engr. Conf., Chicago, IL, Sept. 26–30, 1999, Keynote address.

[19] K. G. Coffman and A. M. Odlyzko, "Growth of the internet," in *Optical Fiber Telecommunications IVB*, I. P. Kaminow and T. Li, Eds. San Diego, CA: Academic, 2002, pp. 17–56.

[20] A. Dwivedi and R. E. Wagner, "Traffic model for USA long-distance optical network," in *Proc. Conf. Opt. Fiber Commun.*, Baltimore, MD, 2000, Paper TuK1, pp. 156–158.

[21] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.



Steven K. Korotky (M'87–SM'87–F'98) received the A.B. degree in physics (graduating Phi Beta Kappa with high honors) from Rutgers University, Piscataway, NJ, in 1975 and the M.S. and Ph.D. degrees in physics from Yale University, New Haven, CT, in 1976 and 1980, respectively.

He is currently Director, Optical Networking Technology and Applications Research, at Bell Laboratories, Lucent Technologies, Holmdel, NJ. Since joining Bell Laboratoriess in 1980, the focus of his research and development activities have been in the areas of integrated-optics and lightwave communications, with emphasis on optical waveguides, high-speed optical modulation and pulse generation, and more recently, optical switching and cross connects, and metro and long-distance photonic transmission systems and networks. He has served two three-year terms as Topical Editor for *Integrated-Optics for the Lasers and Photonics Division of Applied Optics* and has been a Guest Editor for the Optical Society of America (OSA). He holds 18 patents and has coauthored more than 150 journal and conference publications and four books on the subject of integrated optics and optical fiber communication.

Dr. Korotky is a Member of the American Physical Society and a Fellow of the OSA. He received the Bell Laboratories Distinguished Member of Technical Staff Award in 1990 and was corecipient of the 2000 Bell Labs President's Gold Award. He has participated on the program and committees of the Conference on Optical Fiber Communication, the European Conference on Integrated-Optics, the Integrated Photonics Research Conference, and the Micro-Optics Conference. He has also served terms on the Advisory Committee of IPR and the Standing Committee of the ECIO. He has also served as a Guest Editor for the IEEE and is currently serving a term on the IEEE Lasers & Electro-Optics Society (LEOS) Board of Governors as a representative on the JOURNAL OF LIGHTWAVE TECHNOLOGY Steering Committee.