

# Geometric Capacity Provisioning for Wavelength-Switched WDM Networks

Li-Wei Chen, Eytan Modiano

**Abstract**— In this paper, we use an asymptotic analysis similar to the sphere-packing argument in the proof of Shannon’s channel capacity theorem to derive optimal provisioning requirements for networks with both static and dynamic provisioning. We consider an  $N$ -user shared-link model where  $W_s$  wavelengths are statically assigned to each user, and a common pool of  $W_d$  wavelengths are available to all users. We derive the minimum values of  $W_s$  and  $W_d$  required to achieve asymptotically non-blocking performance as the number of users  $N$  becomes large. We also show that it is always optimal to statically provision at least enough wavelengths to support the mean of the traffic.

## I. INTRODUCTION

IN GENERAL, an optical network can consist of a large number of nodes connected in some arbitrary fashion (see Figure 1) and can present the network architect with a complex provisioning problem over multiple links. For simplicity, in this paper we will focus on provisioning a single shared link on a backbone network. Figure 1 also shows a model for the shared colored link in the arbitrary network. We consider provisioning for traffic traveling from left to right along the link. Each wavelength on the link can be used to support one lightpath from one of the incoming fibers on the left side of the link to one of the outgoing fibers on the right side of the link.

Broadly speaking, wavelength provisioning can be done in one of two ways. One option is to *statically* provision a wavelength by hard-wiring the nodes at the ends of the link to always route the wavelength from a given input fiber to a given output fiber. The advantage to this is that the cost of the hardware required to support static provisioning is relatively low: no switching capability or intelligent decision-making ability is required. The downside is a lack of flexibility in using that wavelength – even if the wavelength is not needed to support a lightpath between the assigned input and output fibers, it cannot be

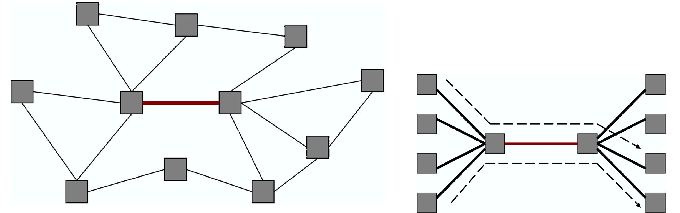


Fig. 1. An example of a mesh optical network consisting of numerous nodes and links, followed by a shared-link model based on the colored link. The dotted lines denote different users of the link. Since each pair of input-output fibers comprises a different user, and there are 4 input fibers and 4 output fibers, there are a total of  $4 \cdot 4 = 16$  users in this example.

assigned to support a lightpath between any other pair of fibers.

This shortcoming can be overcome by using *dynamic* provisioning. A dynamically provisioned wavelength is switched at the nodes on both sides of the link, allowing it to be dynamically assigned to support a lightpath between any source and destination fibers. Furthermore, this assignment can change over time as traffic demands change. This obviously imparts a great deal of additional flexibility. The downside is that the added switching and processing hardware makes it more expensive to dynamically provision wavelengths.

There has been much investigation of both statically provisioned and dynamically provisioned systems in the literature [1], [2], [3], [4]. Such approaches are well-suited for cases where either the traffic is known a priori and can be statically provisioned, or is extremely unpredictable and needs to be dynamically provisioned. However, in practice, due to statistical multiplexing it is common to see traffic demands characterized by a large mean and a small variance around the mean. A hybrid system is well suited to such a scenario. In a hybrid system, a sufficient number of wavelengths are statically provisioned to support the majority of the traffic. Then, on top of this, a smaller number of wavelengths are dynamically provisioned to support the inevitable variation in the realized traffic. Such an approach takes advantage of the relative predictability of the traffic by cheaply provisioning the ma-

The authors are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA (email: lwchen@mit.edu, modiano@mit.edu)

This work was supported by the National Science Foundation (NSF) under Grant ANI-0073730.

majority of the wavelengths, but retains sufficient flexibility through the minority of dynamic wavelengths that significant wavelength overprovisioning is not necessary.

After describing the system model used in this paper, we will use the asymptotic analysis approach from information theory incorporated in the proof of Shannon’s channel capacity theorem [5] to analyze hybrid networks: we allow the number of users to become large, and consider the minimum provisioning in static and dynamic wavelengths necessary to achieve nonblocking performance (i.e., to guarantee that the probability of any call in the snapshot being blocked goes to zero). We will show that it is always optimal to statically provision enough wavelengths to support the traffic mean. We also fully characterize the optimal provisioning strategy for achieving non-blocking performance with minimal wavelength provisioning.

### A. System Model

In the shared link context, we can consider each incoming-outgoing pair of fibers to be a different *user* of the link. Each lightpath request (which we will henceforth term a *call*) can therefore be thought of as belonging to the user corresponding to the incoming-outgoing fiber pair that it uses. We can similarly associate each static wavelength with the corresponding user. Under these definitions, a call belonging to a given user cannot use a static wavelength belonging to a different user – it must either use a static wavelength belonging to its own user, or employ a dynamic wavelength.

Figure 2 gives a flowchart of the decision process for admitting a call. When a user requests a new call setup, the link checks to see if a static wavelength for that user is free. If there is a free static wavelength, it is used. If not, then the link checks to see if any of the shared dynamic wavelengths are free – if so, then a dynamic wavelength is used. If not, then no resources are available to support the call, and it is blocked.

There have been several approaches developed in the literature for blocking probability analysis of such systems under Poisson traffic models [6], including the Equivalent Random Traffic (ERT) model [7], [8], [9] and the Hayward approximation [10]. These approximations, while often able to produce good numerical approximations of blocking probability, are purely numerical in nature and do not provide good intuition for guiding the dimensioning of the wavelengths. Furthermore, they assume that the dynamic wavelengths must be individually switched, and do not consider waveband switching.

In this paper, we adopt a snapshot traffic model that leads to closed-form asymptotic analysis and develop

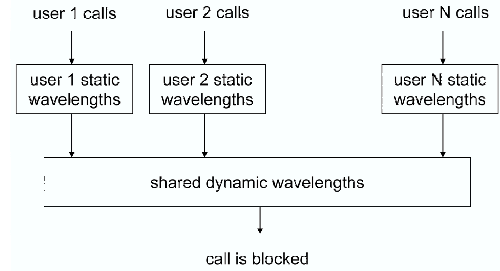


Fig. 2. Decision process for wavelength assignment for a new call arrival. A new call first tries to use a static wavelength if it is available. If not, it tries to use a dynamic wavelength. If again none are available, then it is blocked.

guidelines for efficient dimensioning of hybrid networks. We consider examining a “snapshot” of the traffic demand at some instant in time. The snapshot is composed of the vector  $\mathbf{c} = [c_1, \dots, c_N]$ , where  $c_i$  is the number of calls that user  $i$  has at the instant of the snapshot, and  $N$  is the total number of users.

We model each variable  $c_i$  as a Gaussian random variable with mean  $\mu_i$  and variance  $\sigma_i^2$ . This is reasonable since each “user” actually consists of a collection of source-destination pairs in the larger network that all use the link from the same source fiber to the same destination fiber. In this paper, we will assume that each user has the same mean  $\mu$  and variance  $\sigma^2$ ; the results are extensible to general  $\mu_i$  and  $\sigma_i$  but the extension is beyond the scope of this paper (see [11]). Although the traffic for each individual source-destination pair for the user may have some arbitrary distribution, as long as the distributions are well-behaved, the sum of each traffic stream will appear Gaussian by the Central Limit Theorem.

As a special case, consider the common model of Poisson arrivals and exponential holding times for calls. Then the number of calls that would have entered a non-blocking system at any instant in time is given by the stationary distribution of an  $M/M/\infty$  queue – namely, Poisson with intensity equal to the load  $\rho$  in Erlangs. For a heavy load, this distribution is well approximated by a Gaussian random variable with mean  $\rho$  and variance  $\rho$ .

## II. WAVELENGTH-GRANULARITY SWITCHING

In this section, we consider a shared link, and assume that there are  $N$  users that are the source of calls on the link. Each user is statically provisioned  $W_s$  wavelengths for use **exclusively** by that user. In addition to this static provisioning, we will also provide a total of  $W_d$  dynamically switched wavelengths. These wavelengths can be shared by any of the  $N$  users.

As previously described, we will use a snapshot model of traffic. The traffic is given by a vector  $\mathbf{c} = [c_1, \dots, c_N]$ ,

where each  $c_i$  is independent and identically distributed as  $\mathbf{N}(\mu, \sigma^2)$ . We assume that the mean  $\mu$  is significantly large relative to  $\sigma$  that the probability of “negative traffic” (a physical impossibility) is low, and therefore does not present a significant modeling concern. We will primarily be concerned with a special blocking event that we call *overflow*. An overflow event occurs when there are insufficient resources to support all calls in the snapshot and at least one call is blocked. We will call the probability of this event the *overflow probability*.

From Figure 2, we see that an overflow event occurs if the total number of calls exceeds the ability of the static and dynamic wavelengths to support. This can be expressed mathematically as

$$\sum_{i=1}^N \max \{c_i - W_s, 0\} > W_d \quad (1)$$

where  $\max \{c_i - W_s, 0\}$  is the amount of traffic from each user that exceeds the static provisioning; if the total amount of excess from each user exceeds the available pool of shared dynamic wavelengths, a blocking event occurs.

If we consider the  $N$ -dimensional vector space occupied by  $\mathbf{c}$ , the constraint given by (1) represents a collection of hyperplanes bounding the admissible traffic region:

$$\begin{aligned} c_i &\leq W_s + W_d \\ c_i + c_j &\leq 2W_s + W_d, \quad i \neq j \\ c_i + c_j + c_k &\leq 3W_s + W_d, \quad i \neq j \neq k \\ &\vdots \end{aligned}$$

Each constraint reflect the fact that the sum of the traffic from any subset of users clearly cannot exceed the sum of the static provisioning for those users plus the entire dynamic provisioning available. Note that there are a total of  $N$  sets of constraints, where the  $n^{\text{th}}$  set consists of  $C(N, n) = \frac{N!}{(N-n)!n!}$  equations, each involving the sum of  $n$  elements of the traffic vector  $\mathbf{c}$ . If the traffic snapshot  $\mathbf{c}$  falls within the region defined by the hyperplanes, all calls are admissible; otherwise, an overflow event occurs. The bolded lines in Figure 3 show the admissible region for  $N = 2$  in two dimensions.

#### A. Asymptotic Analysis

We will consider the case where the number of users  $N$  becomes large, and use the law of large numbers to help us draw some conclusions. We can rewrite the call vector in the form

$$\mathbf{c} = \mu \cdot \mathbf{1} + \mathbf{c}'$$

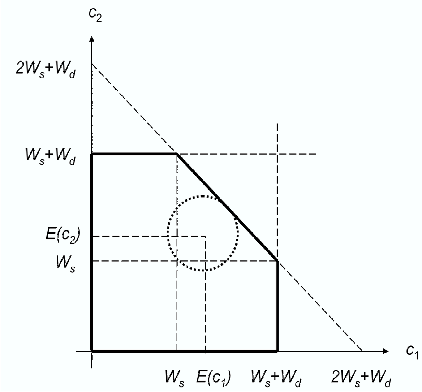


Fig. 3. The admissible traffic region, in two dimensions, for  $N = 2$ . Three lines form the boundary constraints represented by (1). There are two lines each associated with a single element of the call vector  $\mathbf{c}$ , and one line associated with both elements of  $\mathbf{c}$ . The traffic sphere must be entirely contained within this admissible region for the link to be asymptotically non-blocking.

where  $\mathbf{1}$  is the length- $N$  all-ones vector, and  $\mathbf{c}' \sim \mathbf{N}(\mathbf{0}, \sigma^2 \mathbf{1})$  is a zero-mean Gaussian random vector with i.i.d. components. Conceptually, we can visualize the random traffic vector as a random vector  $\mathbf{c}'$  centered at  $\mu \mathbf{1}$ . The length of this random vector is given by

$$\|\mathbf{c}'\| = \sqrt{\sum_{n=1}^N c_n^2}$$

We will use an approach very similar to the sphere packing argument used in the proof of Shannon’s channel capacity theorem in information theory [5]. We will show that asymptotically as the number of users becomes large, the traffic vector falls onto a sphere centered at the mean, and the provisioning becomes a problem of choosing the appropriate number of static and dynamic wavelengths so that this traffic sphere is completely contained within the admissible region.

From the law of large numbers, we know that

$$\frac{1}{N} \sum_{n=1}^N c_n^2 \rightarrow \sigma^2$$

as  $N \rightarrow \infty$ . This implies that asymptotically, as the number of users becomes large, the call vector  $\mathbf{c}$  becomes concentrated on a sphere of radius  $\sqrt{N} \sigma$  centered at the mean  $\mu \mathbf{1}$ . Therefore, in order for the overflow probability to converge to zero, a necessary and sufficient condition is that the hyperplanes described by (1) enclose the sphere entirely. This is illustrated in Figure 3.

#### B. Minimum Distance Constraints

Next, we will derive necessary and sufficient conditions for the admissible traffic region to enclose the traffic

sphere. Our goal is to ensure that we provision  $W_s$  and  $W_d$  such that the minimum distance from the center of the traffic sphere to the boundary of the admissible region is at least the radius of the sphere, therefore ensuring that all the traffic will fall within the admissible region.

Due to the identical distribution of the traffic for each user, the mean point  $\mu\mathbf{1}$  will be equidistant from all planes whose description involves the same number of elements of  $\mathbf{c}$ . We define a *distance function*  $f(n)$  such that  $f(n)$  is the minimum distance from the mean  $\mu\mathbf{1}$  to any hyperplane whose description involves  $n$  components of  $\mathbf{c}$ .

*Lemma 1:* The distance function  $f(n)$  from the traffic mean to a hyperplane involving  $n$  elements of the traffic vector  $\mathbf{c}$  is given by

$$f(n) = \sqrt{n} \left( W_s + \frac{W_d}{n} - \mu \right), \quad n = 1, \dots, N \quad (2)$$

*Proof:* The distance can be calculated using basic geometric principles and is omitted for brevity. ■

We define the *minimum boundary distance* to be

$$F_{min} = \min_{n=1, \dots, N} f(n)$$

A necessary and sufficient condition for the overflow probability to go to zero asymptotically with the number of users is

$$F_{min} \geq \sqrt{N}\sigma$$

We would like to determine the index  $n$  such that  $f(n)$  is minimized. Unfortunately, this value of  $n$  turns out to depend on the choice of provisioning  $W_s$ . Let us consider the derivative of the distance function  $f'(n)$ :

$$f'(n) = \frac{1}{2\sqrt{n}} \left( W_s - \frac{W_d}{n} - \mu \right)$$

We can divide  $W_s$  into three regimes of interest, corresponding to different ranges of values for  $W_s$  and  $W_d$ , and characterize  $f(n)$  in each of these regions:

**Regime 1:** If  $W_s \leq \mu$ :

In this region,  $f'(n) < 0$  for all  $n$ . This implies that  $f(n)$  is a decreasing function of  $n$ , and  $F_{min} = f(N)$ , giving a minimum distance of

$$F_{min} = \sqrt{N} \left( W_s + \frac{W_d}{N} - \mu \right)$$

**Regime 2:** If  $\mu < W_s \leq \mu + W_d$ :

In this region,  $f'(n)$  starts out negative and ends up positive over  $1 \leq n \leq N$ . This implies that  $f(n)$  is convex

and has a minimum. Neglecting integrality concerns, this minimum occurs when  $f'(n) = 0$ , or

$$n^* = \frac{W_d}{W_s - \mu}$$

Therefore  $F_{min} = f(n^*)$  in this regime. Substituting the appropriate values, it can be shown that the minimum distance is given by

$$F_{min} = 2\sqrt{W_d(W_s - \mu)}$$

**Regime 3:** If  $W_s > \mu + W_d$ :

In this region,  $f'(n) > 0$  for all  $n$ . This implies that  $f(n)$  is an increasing function of  $n$ , and  $F_{min} = f(1)$ , giving a minimum distance of

$$F_{min} = W_s + W_d - \mu$$

### C. Optimal Provisioning

In the preceding section, we derived the minimum distance criteria for the hybrid system. Given a fixed number of statically allocated wavelengths  $W_s$ , we can use the equation  $F_{min} \geq \sqrt{N}\sigma$  to calculate the minimum number of dynamic wavelengths  $W_d$  to achieve asymptotically non-overflow performance. We can also draw a few additional conclusions about provisioning hybrid systems.

*Theorem 1:* A minimum of  $\mu$  static wavelengths should always be provisioned per user.

*Proof:* For  $W_s \leq \mu$ , we know from Case 1 above that the minimum distance constraint is

$$\begin{aligned} F_{min} &= \sqrt{N} \left( W_s + \frac{W_d}{N} - \mu \right) \geq \sqrt{N}\sigma \\ \Rightarrow W_{tot} &= NW_s + W_d \geq (\mu + \sigma)N \end{aligned}$$

Note that the total number of wavelengths  $W_{tot} = NW_s + W_d$  is independent of  $W_s$  and  $W_d$  in this regime, suggesting that the same total number of wavelengths are required regardless of the partitioning between static and dynamic wavelengths. Since static wavelengths are less expensive to provision than dynamic wavelengths, this shows that there is never any reason to provision less than  $W_s = \mu$  wavelengths. ■

An interesting corollary to this theorem follows from the observation that the case where  $W_s = 0$  (i.e. all wavelengths are dynamic) also falls in this regime (i.e. Regime 1). Since fully dynamic provisioning is obviously the least-constrained version of this system, we can use it as a bound on the minimum number of wavelengths required by **any** asymptotically overflow-free system:

$$W_{tot} \geq (\mu + \sigma)N$$

We can also consider a system that is fully static, with no dynamic provisioning. This is the most inflexible wavelength partitioning, and provides us with an upper bound on the number of wavelengths required by any hybrid system.

*Theorem 2:* For a fully static system with no dynamic provisioning, the minimum number of wavelengths required is given by

$$W_{tot} = (\mu + \sigma)N + (\sqrt{N} - 1)N\sigma$$

*Proof:* Let  $W_d = 0$ . Then, for overflow-free operation, we obviously need  $W_s > \mu$ . This puts us in Regime 3 where  $W_s > \mu + W_d$ , and the minimum distance condition gives us

$$\begin{aligned} F_{min} &= W_s + W_d - \mu > \sqrt{N}\sigma \\ \Rightarrow W_{tot} &= (\mu + \sigma)N + (\sqrt{N} - 1)N\sigma \end{aligned}$$

Note that this exceeds the lower bound on the minimum number of wavelengths by  $(\sqrt{N} - 1)N\sigma$ . We can therefore regard this quantity as the **maximum switching gain** that we can achieve in the hybrid system. This gain is measured in the maximum number of wavelengths that could be saved if all wavelengths were dynamically switched.

*Corollary:* Combining the upper and lower bounds, we observe that for efficient overflow-free operation, the total number of wavelengths required by any hybrid system is bounded by

$$(\mu + \sigma)N \leq W_{tot} \leq (\mu + \sigma)N + (\sqrt{N} - 1)N\sigma$$

#### D. Numerical Example

Simulations were conducted to verify the accuracy of the provisioning results derived. Figure 4 verifies the results of the preceding discussion for the case of  $\mu = 100$  and  $\sigma = 10$ . The rapidly descending curve shows that if the theoretical minimum of  $W_{tot} = (\mu + \sigma)N$  wavelengths are provisioned with  $W_s = \mu$ , then as  $N$  increases, the overflow probability drops off quickly and eventually the system becomes asymptotically non-blocking. The other two curves show that if less than  $W_{tot}$  wavelengths are provisioned, the overflow probability no longer converges to zero as the number of users increases.

Note also that the convergence occurs fairly rapidly if the  $W_{tot}$  wavelengths calculated in the preceding sections are provisioned. In a system with just 30 users, the

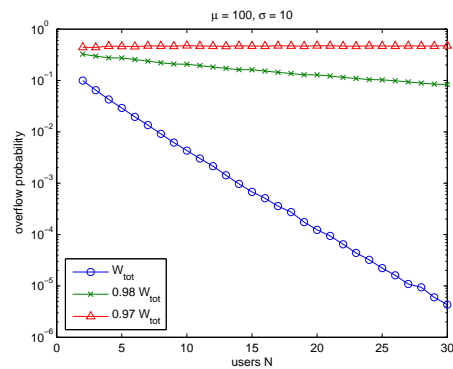


Fig. 4. Curves show decrease in overflow probability with increasing number of users  $N$ . The curve with the circles shows a link provisioned with the theoretical minimum number of wavelengths  $W_{tot}$  needed to achieve asymptotically non-overflowing operation. Note that if fewer than  $W_{tot}$  wavelengths are provisioned, the overflow probability no longer converges to zero as the number of users increases.

overflow probability has already decreased to the order of  $10^{-5}$ . Since the number of users is equal to the number of input-output fiber pairs, this corresponds to a link with as few as 5 input fibers and 6 output fibers, for example. Therefore, the results are useful in designing for good network performance even when  $N$  is finite and small.

#### REFERENCES

- [1] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*, Morgan Kaufmann, 1998.
- [2] L. Li and A. K. Somani, "Dynamic wavelength routing using congestion and neighborhood information," *IEEE/ACM Trans. Networking*, vol. 7, pp. 779–786, October 1999.
- [3] A. Birman, "Computing approximate blocking probabilities for a class of all-optical networks," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 852–857, June 1996.
- [4] O. Gerstel, G. Sasaki, S. Kutten, and R. Ramaswami, "Worst-case analysis of dynamic wavelength allocation in optical networks," *IEEE/ACM Trans. Networking*, vol. 7, pp. 833–845, December 1999.
- [5] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley-Interscience, 1991.
- [6] R. Guerin and L. Y.-C. Lien, "Overflow analysis for finite waiting-room systems," *IEEE Trans. Commun.*, vol. 38, pp. 1569–1577, September 1990.
- [7] R. I. Wilkinson, "Theories of toll traffic engineering in the u.s.a.," *Bell Syst. Tech. J.*, vol. 35, pp. 412–514, March 1956.
- [8] R. B. Cooper, *Introduction to Queueing Theory, 2nd Ed.*, North Holland, New York, 1981.
- [9] D. A. Garbin M. J. Fischer and G. W. Swinsky, "An enhanced extension to wilkinson's equivalent random technique with application to traffic engineering," *IEEE Trans. Commun.*, vol. 32, pp. 1–4, January 1984.
- [10] A. A. Fredericks, "Congestion in blocking systems - a simple approximation technique," *Bell Syst. Tech. J.*, vol. 59, pp. 805–827, July-August 1980.
- [11] L. Chen, *A Study on the Tradeoff Between Efficient Resource Allocation and Node Complexity in WDM Networks*, Ph.D. thesis, MIT, September 2005.