Circuit Allocation in All Optical Networks with Average Packet Delay Cost Criterion[1]

Zvi Rosberg[2]

Abstract

An optimal circuit allocation problem for all-optical circuit-switched backbone networks with average packet delay cost criterion is considered. Multiple classes of traffic flows arrive at the network edge routers, where they can be queued until an end-to-end optical circuit can be allocated. Assuming fluid traffic and circuit allocation of fixed periods, a lower bound on the optimal average packet delay is derived and the cost of two allocation policies are exactly evaluated. The cost of both policies are demonstrated for a variety of deterministic and random networks and are compared with the lower bound.

Index Terms

Optical networks, Circuit switching, Golden ratio, Markov decision processes, Packet delay, TDM.

I. INTRODUCTION

Circuit switching is a classical network architecture used for real time applications such as telephony, which is also considered attractive for all-optical backbone networks (AON) [3] [14] [15]. Circuit switching in general, provides guaranteed bandwidth and low delay. Optical circuit switching also avoids the electronics associated with high-speed queuing and scheduling hardware at the core routers. Understanding the potential merit of circuit switching for optical networks is becoming of utmost important in light of hybrid switching methods [7] [18] comprising of optical burst switching (OBS) [2] and optical circuit switching (OCS) [6] [8].

---

[2]CSE Dept., Ben Gurion University, Beer-Sheva, 84105, Israel. Email: rosberg@bgu.ac.il

AON comprises buffered edge routers at the ingress/egress network nodes and bufferless core routers inside the network. Edge routers multiplex/demultiplex end user traffic flows, e.g., SONET [1], and due to bufferless core routers apply a reservation protocol before transmitting buffered data into the network.

With circuit switching, a two-way reservation protocol is used, by which circuits are allocated for a period of time allowing lossless traffic flows from a subset of sources to their corresponding destinations. Using wavelength division multiplexing (WDM), an optical circuit comprises switching resources and a sequence of wavelengths along a multiple-hop route conforming the wavelength conversion rules [17]. An optical circuit corresponds to a unidirectional lightpath between a source and destination pair of edge routers.

A fundamental control problem in a circuit-switched AON is to find an adaptive policy based on the buffer occupancies to allocate circuits (lightpaths) so as to minimize the average packet delay. Note that with circuit switching, queuing delay is the main concern since by proper edge buffer sizing, packet loss diminishes.

This paper derives a lower bound to the optimal long-run average packet delay and evaluates the performance of two heuristic policies. Previous studies have analyzed circuit-switched networks with respect to blocking probability, or equivalently carried traffic. The study in [10] concerns with routing data or voice in a classical circuit-switched network and the studies in [13] [15] concern with *Routing and Wavelength Assignment (RWA)* in optical circuit-switched networks. The blocking probabilities have been evaluated there by using reduced-load fixed point approximation. In [4] [15], carried traffic maximization has been formulated as an integer linear program (ILP). Although ILP is in NP-hard, its solution can be derived off-line for each traffic demand, after which it can be applied on-line using a lookup table.

A framework for evaluating packet delay in AON with arbitrary adaptive RWA algorithms and two-way reservation policies is derived in a related study [16]. The framework there is applicable to any stationary RWA policy but not to policies that multiplex circuit subsets over the time domain, i.e., *TDM policies*. Such policies are used in Subsection V-C to approximate the optimal policy.

In Section II, the system model is defined and the optimization problem is formulated. In Section III, a sufficient and necessary condition for the existence of a finite valued optimal allocation policy is derived. The optimal cost is bounded from below in Section IV and two

allocation policies are proposed and analyzed in Section V. Their costs and the lower bound in a variety of deterministic and random networks are demonstrated in Section VI and a brief conclusion is given in Section VII.

II. MODEL FORMULATION

A traffic flow is characterized by a source-destination pair of edge routers, QoS class and other external classifications. Packets that cannot be transmitted immediately are queued in an unbounded logical buffer associated with that stream. Buffers are labeled by j, where $1 \leq j \leq J$.

A *circuit* is a unidirectional lightpath connecting a source-destination pair of edge routers capable of uninterruptedly transmitting C b/s for an integral period of T seconds. A circuit is setup by selecting a feasible path and reserving a feasible sequence of wavelengths and resources along that path. An allocated circuit drains its corresponding buffer at a maximum rate of C b/s. After T seconds, the circuit can be teared down.

Every T seconds, a set of circuits, referred to as a *transmission set*, is selected and allocated to their corresponding buffers using an arbitrary selection algorithm. The transmission set can be selected as a function of all logical queue lengths.

Two circuits are *wavelength-disjoint*, if their wavelengths in each fiber are different. Allocating a transmission set containing non wavelength-disjoint circuits is not optimal since part of the traffic is lost. Thus, by excluding one of the conflicting circuits, carried traffic is increased and results in a lower value of the cost function. The relation between the carried traffic and the cost function is made clear in Section III. A *transmission set is maximal*, if it cannot be extended with another wavelength-disjoint circuit. Allocating a non-maximal transmission set is also not optimal since by extending it to a maximal set, traffic is increased.

Circuit allocation comprises two tasks: (i) Specification of the maximal transmission sets; and (ii) Selection of a maximal transmission set every T seconds. The first task is the classical combinatorial problem known as *disjoint paths problems* [11], which is in NP-hard [9]. Task (i) is pursued only once by representing the network topology as a graph, and then finding all maximal transmission sets. It is complex and heuristic shortcuts can be done during the graph representation part, some of which are to constrain the lightpath hop-count and to limit the number of lightpaths connecting each source-destination pair. Disjoint paths problems have been extensively studied in [9] and in references there. This paper focuses on task (ii), where a given

group of maximal transmission sets are adaptively scheduled over the time domain.

Since optical links have a capacity of 1G-1T b/s and optical switching is on-the-fly, fluid flows is a conventional way to specify user traffic in core AON. Let A_j be the integral valued rate of b/s arriving at the logical buffer j.

For scalability reduction, the circuit period and rates are normalized as follows without losing generality. Time scale is measured in circuit periods, i.e., $T = 1$ is assumed and transmission and arrival rates are specified in circuit periods. Furthermore, rates are normalized by dividing them with their largest common (integral) denominator. Normalized rates are referred to as rates in *B-bits*. After normalization, K and $\{A_j, 1 \leq j \leq J\}$ denote the lightpath transmission and arrival rates, respectively, in B-bits per circuit period. That is, time units are specified in circuit periods and data units are specified in B-bits.

The circuit allocation problem is formulated by the following Markov Decision Process (MDP). Let $\mathcal{J}_i \subseteq \{1, 2, \ldots, J\}$, $1 \leq i \leq N$, be the family of all maximal transmission sets; $X_j(n)$ be the queue length (in B-bits) of buffer j at step n; and $\delta_i(n)$ be 1 or 0, depending on whether or not transmission set i is selected at step n, respectively. Note that the number of traffic sources is J and the number of maximal transmission set is N.

Without losing optimality, only policies that select exactly one maximal transmission set at each step can be considered. Otherwise, part of the carried traffic is lost. Thus, $\sum_{i=1}^{N} \delta_i(n) = 1$, for every n. By the model definition, the evolution of $X_j(n)$, for every j, is given by

$$X_j(n + 1) = \left[X_j(n) + A_j - K \sum_{i=1}^{N} \delta_i(n) I_{\{j \in \mathcal{J}_i\}} \right]^+, \tag{1}$$

where $[y]^+ = \max\{0, y\}$ and $I_{\{j \in \mathcal{J}_i\}}$ is 1 or 0 depending on whether or not $j \in \mathcal{J}_i$, respectively. That is,

$$X_j(n + 1) = \begin{cases} [X_j(n) + A_j - K]^+, & \text{if } j \text{ is selected,} \\ X_j(n) + A_j, & \text{Otherwise.} \end{cases} \tag{2}$$

The *immediate cost* at step n, $C(n)$, is defined as the time-average of all queue lengths from the beginning of step n until the beginning of step $n + 1$. That is, if transmission set $\mathcal{J}_i$ is selected at step n (i.e., $\delta_i(n) = 1$), then

$$\begin{aligned} C(n) &= \sum_{j \in \mathcal{J}_i} \int_{t=0}^{1} \left[X_j(n) + A_j t - Kt \right]^+ dt + \sum_{j \notin \mathcal{J}_i} \int_{t=0}^{1} (X_j(n) + A_j t) dt \\ &\sum_{j \in \mathcal{J}_i} \left(\frac{[\min\{K - A_j, X_j(n)\}]^2}{2(K - A_j)} + [X_j(n) - (K - A_j)]^+ \right) + \sum_{j \notin \mathcal{J}_i} \left(X_j(n) + \frac{A_j}{2} \right). \end{aligned} \tag{3}$$

The last equality follows from the fact that arrival and service rates are fixed and the step length is normalized to one.

A circuit allocation policy π selects a transmission set at every circuit allocation step n. Generally, π may take random selections depending on the step n and the history of queue lengths. For every policy π and every initial state $\boldsymbol{x}$ (i.e., $\boldsymbol{X}(0) = \boldsymbol{x}$), let $V_n(\pi, \boldsymbol{x})$ be the total expected cost during the first n steps. The long-run average cost of policy π is defined by

$$\overline{V}(\pi, \boldsymbol{x}) = \limsup_{n \to \infty} \frac{V_n(\pi, \boldsymbol{x})}{n} \ . \tag{4}$$

For every stationary policy π whose underlying Markov process is ergodic, $\overline{V}(\pi, \boldsymbol{x})$ is the expected total queue lengths under stationary conditions. Furthermore, if the lengths of the packets arriving at buffer j are independent and identically distributed random variables with mean M_j, Little's lemma implies that $\overline{V}(\pi, \boldsymbol{x})/\sum_{j=1}^{J} M_j$ is the expected delay of an arbitrary packet under stationary conditions.

A policy π^* is optimal, if it attains the *value function* $\inf_\pi \overline{V}(\pi, \boldsymbol{x})$ for every initial state $\boldsymbol{x}$. A necessary and sufficient condition for a finite value function is derived in the next section. Also, since the mean packet lengths $\{M_j\}$ are given in the outset, the same policy minimizes both, the long-run total average queue lengths and the long-run average packet delay.

It will become apparent in Section IV that the alternative 'on-off' arrival process (rather than the current 'always-on') would not change the lower bound on the value function. Moreover, it could only marginally effect the cost of TDM policies analyzed in Section V since the accumulated arrivals between two consecutive circuit allocations is dominated by the average arrival rate.

III. A CONDITION FOR A FINITE VALUE FUNCTION

Since bits arrive at constant rates independently of the allocation policy, the value function in (4) is attained by a policy maximizing the long-run average of the carried traffic rate. Consequently, as explained in Section II, attention can be restricted to a policy that for at each step n selects exactly one maximal transmission set. Otherwise the policy can be improved by another policy with a larger carried traffic rate.

The following theorem provides a necessary and sufficient condition for the existence of a finite value function.

Theorem 1: The value function is finite if and only if there is a set of proportions α_i, $1 \leq i \leq N$, such that

$$\sum_{i=1}^{N} \alpha_i \leq 1 \ \ \text{and} \ \ \rho_j \stackrel{\text{def}}{=} \sum_{i:j \in \mathcal{J}_i} \alpha_i \geq \frac{A_j}{K}, \ \forall \ 1 \leq j \leq J. \tag{5}$$

Proof: For sufficiency it will be shown that (5) implies a policy π_0 with a finite long-run average total queue length.

Since K and $\{A_j\}$ are integral numbers, $\{A_j/K\}$ are rationales. Since rationales are dense in the reals, (5) implies the existence of rationales $\{\tilde{\rho}_j, \tilde{\rho}_j \leq \rho_j\}$, satisfying (5). Since $\tilde{\rho}_j \leq \rho_j$, $\forall \ j$, there are also rationales $\{\tilde{\alpha}_i\}_{i=1}^{N}$ satisfying (5).

Policy π_0 is defined as follows. Let m be a common denominator of $\{\tilde{\alpha}_i\}$. For the first m steps, π_0 does nothing. For every subsequent cycle of m steps, π_0 selects each transmission set $\mathcal{J}_i$, for $m_i = m\tilde{\alpha}_i$ consecutive steps according to the set label order. Thus, in every cycle of m steps each buffer j transmits during $m\tilde{\rho}_j$ steps.

For any given $\boldsymbol{X}(0) = \boldsymbol{x}$, $\boldsymbol{X}(m)$ under policy π_0 is $\boldsymbol{X}(0) + (mA_1, mA_2, \ldots mA_J)$. Combined with (5), it follows that $\{\boldsymbol{X}(lm) \mid l \geq 1\}$ is a non-increasing (component-wise) sequence and straightforward calculation reveals that the long-run average of the total queue length is finite.

The necessity of condition (5) is shown by a contradiction argument. Suppose that the value function is finite and is attained by policy π^*. Let $\alpha_i(\pi^*)$ be the long-run proportion that transmission set $\mathcal{J}_i$ is selected by policy π^* and define $\rho_j(\pi^*) = \sum_{i:j \in \mathcal{J}_i} \alpha_i(\pi^*)$, $1 \leq j \leq J$.

Suppose in contradiction that there is a buffer j_0 for which

$$\rho_{j_0}(\pi^*) < A_{j_0}/K - \epsilon, \quad \text{for some } \epsilon > 0. \tag{6}$$

Let $n_{j_0}(n, \pi^*)$ be the number of times policy π^* allocates a circuit to buffer j_0 during the first n steps. Since the long-run proportions $\{\alpha_i(\pi^*)\}$ exist, it follows by the evolution in (2) that for sufficiently large n,

$$n_{j_0}(n, \pi^*)/n < \rho_{j_0}(\pi^*) + \epsilon/2. \tag{7}$$

Thus, by (6)–(7), for every initial state $\boldsymbol{x}$ and every step n,

$$\begin{aligned} X_{j_0}(n) \ &\geq X_{j_0}(0) + nA_{j_0} - n_{j_0}(n, \pi^*)K = \left(\tfrac{X_{j_0}(0)}{n} + A_{j_0} - \tfrac{n_{j_0}(n,\pi^*)}{n}K \right) n \\ &\geq \left(A_{j_0} - \rho_{j_0}(\pi^*)K - \tfrac{\epsilon K}{2} \right) n > \left(\epsilon K - \tfrac{\epsilon K}{2} \right) n = \tfrac{n\epsilon K}{2} \ . \end{aligned} \tag{8}$$

By (8), $\lim_{n \to \infty} X_{j_0}(n) = \infty$, which is in contradiction with the finiteness of $\overline{V}(\pi, \boldsymbol{x})$. ∎

Note that the necessity of (5) implies that attention can be restricted to policies π satisfying

$$\rho_j(\pi) \geq \frac{A_j}{K} , \quad \text{for every } 1 \leq j \leq J . \tag{9}$$

These policies are referred to as *feasible policies*.

IV. A LOWER BOUND TO THE VALUE FUNCTION

Finding a closed form representation of the optimal policy is intractable. Therefore, the value function is bounded by a lower bound which is used as a reference for heuristic policies.

Every feasible policy, π, is associated with a feasible proportion vector in the set $\mathcal{P} = \{\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)\}$, where each $\boldsymbol{\alpha}$ satisfies (5).

A. Constrained lower bounds on individual queues

For feasible allocation proportion ρ_j (determined by $\boldsymbol{\alpha} \in \mathcal{P}$), let $\Pi(\rho_j)$ be the set of feasible policies π with $\rho_j(\pi) = \rho_j$. First, the long-run average queue length in buffer j is bounded below for all $\pi \in \Pi(\rho_j)$.

For every feasible policy π, (2) implies that the long-run average queue length in buffer j, $E_\pi(X_j)$, can be expressed by the following function $S_Y^j(D)$. For every buffer j and non-negative integers Y and D, define

$$S_Y^j(D) = \frac{Y \cdot (Y - [Y + A_j - K]^+)}{2\max\{Y, K - A_j\}} + D \cdot [Y + A_j - K]^+ + \frac{A_j(D-1)^2}{2}, \tag{10}$$

where Y represents the queue length at the beginning of a circuit allocation period and D represents the number of circuit periods until the next allocation (excluding the current allocation period but including the next allocation period). Note that given Y and D, $S_Y^j(D)$ is the integral of the queue length in buffer j between two consecutive allocations.

Also, every stationary $\pi \in \Pi(\rho_j)$ induces a deterministic sequence $\{(D_t, Y(t)) : t = 1, 2, \ldots\}$, where D_t is the number of circuit periods between allocation t (excluded) and allocation $t + 1$ (included) and $Y(t)$ is the queue length in buffer j when policy π allocates the t-th circuit.

Since ρ_j is the long run proportion of circuits allocated to buffer j, it follows that

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} D_t = \frac{1}{\rho_j}. \tag{11}$$

By (2) and (11),

$$E_\pi\left(X_j\right) = \frac{\lim\limits_{T\to\infty}\sum\limits_{t=1}^{T}S^j_{Y(t)}(D_t)}{\lim\limits_{T\to\infty}\sum\limits_{t=1}^{T}D_t} = \rho_j \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}S^j_{Y(t)}(D_t). \tag{12}$$

By (10), the r.h.s of (12) is convex in $\{D_t\}$. Thus, recalling that $\{(D_t, Y(t))\}$ is a deterministic sequence, Jensen's inequality, (11) and (12) imply

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}S^j_{Y(t)}(D_t) \geq \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}S^j_{Y(t)}\left(\frac{1}{\rho_j}\right) = S^j_{A_j(1-\rho_j)/\rho_j}(1/\rho_j), \tag{13}$$

where the last equality follows from the fact that $A_j(1-\rho_j)/\rho_j$ is the stationary queue length at a circuit allocation instant given a constant inter-allocation period of length $1/\rho_j$.

Combing (10), (12) and (14) yields

$$E_\pi\left(X_j\right) \geq \frac{(1-\rho_j)^2 A_j K}{2\rho_j(K-A_j)}. \tag{14}$$

B. A convex program

Since $\overline{V}(\pi, \boldsymbol{x}) = \sum_{j=1}^{J}E_\pi\left(X_j \mid X_j(0)=x_j\right)$, (14) implies that for every feasible $(\rho_1, \rho_2, \ldots, \rho_J)$,

$$\overline{V}(\pi, \boldsymbol{x}) \geq \sum_{j=1}^{J}\frac{(1-\rho_j)^2 A_j K}{2\rho_j(K-A_j)}, \qquad \forall\ \pi \in \bigcap_{j=1}^{J}\Pi(\rho_j). \tag{15}$$

Moreover, every feasible π resides in some set $\bigcap_{j=1}^{J}\Pi(\rho_j)$, where the proportions $\{\rho_j\}$ satisfy condition (5). Thus, the value function is bounded below by the solution to

$$\min_{\rho_1,\ldots,\rho_J}\sum_{j=1}^{J}\frac{(1-\rho_j)^2 A_j K}{2\rho_j(K-A_j)} \tag{16}$$

$$\text{subject to}:\ \ \rho_j = \sum_{i:j\in\mathcal{J}_i}\alpha_i \geq \frac{A_j}{K}\ ,\qquad 1 \leq j \leq J,$$
$$\sum_{i=1}^{N}\alpha_i = 1,\ \ \alpha_i \geq 0,\ \ 1 \leq i \leq N. \tag{17}$$

By replacing every ρ_j in the objective function (16) with its definition in (17), the convex program translates into

$$\min_{\alpha_1,\ldots,\alpha_N}\sum_{j=1}^{J}\frac{A_j K(1-\sum_{i:j\in\mathcal{J}_i}\alpha_i)^2}{2(K-A_j)\sum_{i:j\in\mathcal{J}_i}\alpha_i} \tag{18}$$

$$\text{subject to}:\ \ \sum_{i=1}^{N}\alpha_i = 1,\ \ \alpha_i \geq 0\ ,\ \ 1 \leq i \leq N,$$
$$\sum_{i:j\in\mathcal{J}_i}\alpha_i \geq \frac{A_j}{K},\ \ 1 \leq j \leq J. \tag{19}$$

The convex program (18)–(19) has a unique minimum that can be derived by standard algorithms, e.g., gradient projection, active set and primal method [12]. Mathematical software tools are also available, one of which is applied in Section VI. The solution of (18)–(19) yields a lower bound and ideal allocation proportions for the sets $\{\mathcal{J}_i\}$. Also, the ideal proportions and uniform inter-allocation periods providing the lower bound may serve as a beacon to which network topology and circuit allocation algorithms are aiming at.

Two heuristic policies attempting to approximate the ideal proportions and to distribute the allocation periods of each transmission set as uniform as possible are defined next.

V. TDM POLICIES

For symmetric networks and uniform traffic, the lower bound derived in Section IV is attained by the round-robin allocation policy, where circuits are allocated to transmission sets in a cyclic order. For most of the other cases, the lower bound cannot be attained since uniform inter-circuit-allocation periods across all buffers is infeasible. In this section, two policies approximating uniform inter-allocations are analyzed. One policy, which is referred to as *Longest Time to Empty Queue (LTEQ)*, allocates circuits to transmission sets based on a state-dependent priority index. Another policy allocates circuits to transmission sets in a weighted round-robin fashion according to an order determined by the *golden ratio*.

A. *The LTEQ policy*

LTEQ policy is motivated by simplicity and by many empirical observations indicating that short term optimization usually works well. By (4), the naive one-step optimization rule (myopic) is to allocate the circuits to the transmission set which maximizes the transmission rate at the next step. That is, at every state x, select transmission set

$$\arg \max_{\{\mathcal{J}_i\}} \sum_{j \in \mathcal{J}_i} \min\{x_j + A_j, K\}.$$

Since transmission rate is insensitive to large queue lengths, this policy favors large transmission sets which may result in starvation. A more sensible policy is to select the transmission set with the longest time to empty all its queues given that it is being served without interruption. Specifically, given that transmission set $\mathcal{J}_i$ is assigned consecutive circuits at state x, the time to empty all its queues is $t(x, \mathcal{J}_i) = \max_{j \in \mathcal{J}_i} x_j/(K - A_j)$.

Thus, *at every state x, LTEQ selects the set $\mathcal{J}(x)$ with lowest label which attains*

$$\arg\max_{\{\mathcal{J}_i\}} t(\boldsymbol{x}, \mathcal{J}_i).$$

Note that for arrival streams with fixed rates, the queue lengths with policy *LTEQ* evolves deterministically. Thus, for every given initial state at time 0, the circuit periods allocated to each transmission set, $\mathcal{J}_i$, are completely determined at time 0. Namely, *LTEQ* is a TDM policy.

B. The golden ratio policy

The *Golden Ratio (GR)* policy is a generalization of the symmetric round-robin policy first presented in [5] for a multiple access channel.

Let $\boldsymbol{\alpha}^*$ be the ideal circuit allocation proportions obtained from the solution of (18)–(19). Since the lower bound is attained when the inter-circuit-allocation periods for every buffer j is uniform, one wish to approximate this uniformity without wasting circuit periods. The best known approximation is the following rule based on the golden ratio (defined by $\phi^{-1} = (\sqrt{5} - 1)/2$).

To attain the optimal allocation proportions, $\boldsymbol{\alpha}^*$, with circuit period units, integers τ and $\{\tau_i \mid 1 \leq i \leq N\}$ are fixed first so as to satisfy $\tau_i/\tau \approx \alpha_i^*$, $\forall\, i$, and $\sum_{i=1}^{N} \tau_i = \tau$. (Fibonacci numbers yield the best results [5]). Then, the following *GR* rule is used to order between transmission set allocations within each cycle of τ circuit periods.

Place the τ fractions $a_j = j\phi^{-1} - \lfloor j\phi^{-1} \rfloor$, $j = 0, 1, \ldots, \tau - 1$, on a circle with contour one. Then associate the first τ_1 fractions, $\{a_j, j = 1, \ldots, \tau_1\}$, with transmission set 1, the next τ_2 fractions, $\{a_j, j = \tau_1 + 1, \ldots, \tau_2\}$, with transmission set 2, etc. The resulting order determines the GR allocation order.

C. The long-run average cost of TDM policies

A finite numerical procedure is derived next for exact evaluation of the cost of any TDM policy with a finite cycle. A TDM policy with a finite cycle is specified by a cycle length τ and N integer-valued inter-allocation times, $\boldsymbol{D}_i = (D_i^1, D_i^2, \ldots, D_i^{\tau_i})$ between two consecutive allocations to transmission set $\mathcal{J}_i$ satisfying $\sum_{m=1}^{\tau_i} D_i^m = \tau$, $1 \leq i \leq N$.

Every set of inter-allocation times to transmission sets, $\{\boldsymbol{D}_i \mid 1 \leq i \leq N\}$, induces a set of inter-allocation times to buffers. Let $\{\boldsymbol{d}_j \mid 1 \leq j \leq J\}$ be the induced set, where $\boldsymbol{d}_j = (d_j^1, d_j^2, \ldots, d_j^{t_j})$; d_j^l is the l-th inter-allocation time to buffer j; and t_j is the number of inter-allocation times to buffer j in each cycle.

The evaluation of the total average queue lengths is based on two properties: (i) K and $\{A_j \mid 1 \leq j \leq J\}$ are integers; and (ii) For every TDM policy, each queue length evolves independently of the others.

The non-transient states of a TDM policy with a cycle of length τ and long-run buffer allocation proportions satisfying condition (5) are contained by the interval $[0, \tau A_j]$. Moreover, the feasible states assume only rational values n/m, where m is the common denominator of the integers $K, A_j, 1 \leq j \leq J$. Thus, each process $\{X_j(n) \mid n \geq 1\}$ assumes only a finite number of values and so is the multi-dimensional process $\{\boldsymbol{X}(n) \mid n \geq 1\}$. Combining with the fact that the evolution of $\{X_j(n) \mid n \geq 1\}$ is deterministic, the exact long-run average total queue lengths can be computed by the following finite and very fast procedure.

For each TDM policy, the process $\{\boldsymbol{X}(n) \mid n \geq 1\}$ is examined at the beginning of circuit periods $n = k \cdot \tau$, where $k = 0, 1, 2, \ldots$, and a recurrent state is identified. The existence of a recurrent state is guaranteed by the fact that the set of feasible states under each policy is finite and the decisions are non-randomized and stationary.

Let n_0 and n_1 be the first two circuit periods where a state recurs under policy $\pi \in \{GR, LTEQ\}$. From (2) and the mean ergodic theorem,

$$
\begin{aligned}
\overline{V}(\pi, \boldsymbol{x}) \;\; &= \sum_{j=1}^{J} E\big(X_j \mid X_j(0) = x_j\big) = \frac{1}{n_1 - n_0} \sum_{n=n_0}^{n_1-1} \sum_{i=1}^{N} \delta_i(\pi, n) \\
&\cdot \left(\sum_{j \in \mathcal{J}_i} S^{j}_{X_j(n)}(1) + \sum_{j \notin \mathcal{J}_i} \big(X_j(n) + A_j/2\big) \right),
\end{aligned}
\tag{20}
$$

where $S^{j}_{X_j(n)}(1)$ is given by (10) and $\delta_i(\pi, n)$ is 1 or 0, depending on whether or not transmission set i is selected by policy π at step n, respectively.

The computational procedure in (20) is applied in the next section for evaluating the costs of the *LTEQ* and *GR* policies. Observe that this procedure is exact and **is not evaluation by simulation**, although it emulates the underlying deterministic process.

VI. NUMERICAL EXAMPLES

The quality of the lower bound and the performance of the *GR* and the *LTEQ* policies are demonstrated in deterministic and random networks. Note that the network topology and the algorithm by which the maximal transmission sets are derived are immaterial - only the transmission set composition.

In all the cases, the number of buffers is 100 and the number of transmission sets is 10 and the cases are classified according to the buffer distribution over the transmission sets. For symmetric cases, where all transmission sets comprise the same number of buffers and the number of transmission sets containing any given buffer is the same for every buffer, it is straightforward to prove that both policies attain the lower bound and therefore are optimal. The cases below concern with versatile asymmetric networks.

Asymmetric networks are characterized by two attributes: (i) the distribution of the transmission set size; and (ii) the distribution of the number of transmission sets containing a given buffer. In addition, buffer can be distributed deterministically or randomly. A random buffer distribution reflects an average performance measure over many networks having the same mean characteristics with respect to attributes (i) and (ii). The specific numerical examples are as follows: (D1) Deterministic buffer distribution where attribute (i) is variable and attribute (ii) is fixed; (D2) Deterministic buffer distribution where attribute (i) and (ii) are both variables; (R1) Random buffer distribution where attribute (i) and (ii) are both variables so as to form a random network with symmetric expected values; and (R2) A random asymmetric network.
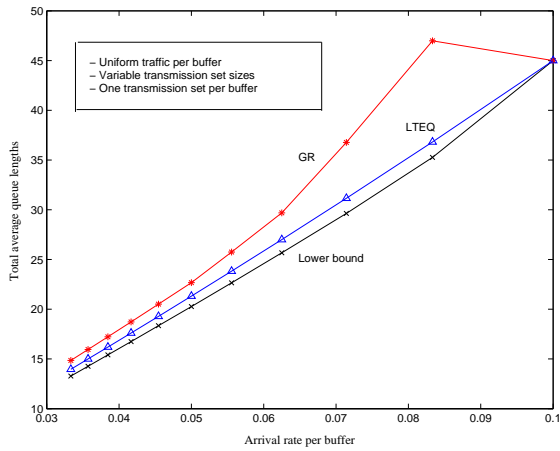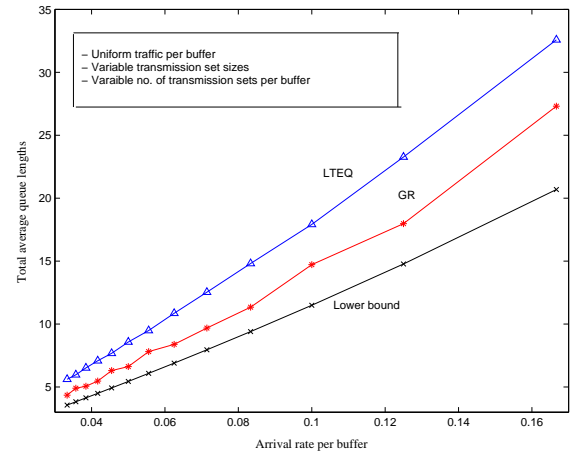

Fig. 1.   Case D1.

Fig. 2.   Case D2.

Typical lower bounds and the costs of both policies for symmetric offered load per buffer ranging from light to heavy until network saturation are depicted in Figures 1–4. The conclusion below are based on many other cases not presented here, where it has been also observed that relative measures do not change significantly when offered loads are asymmetric. For the random
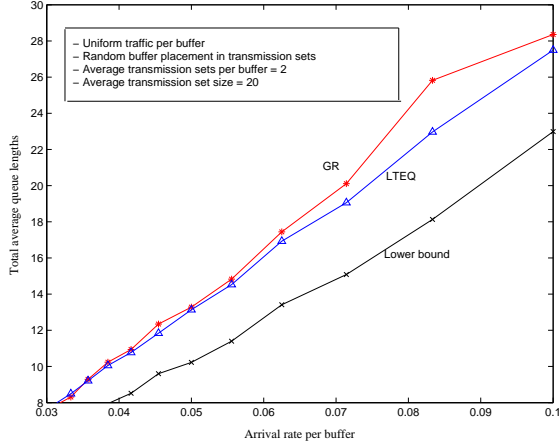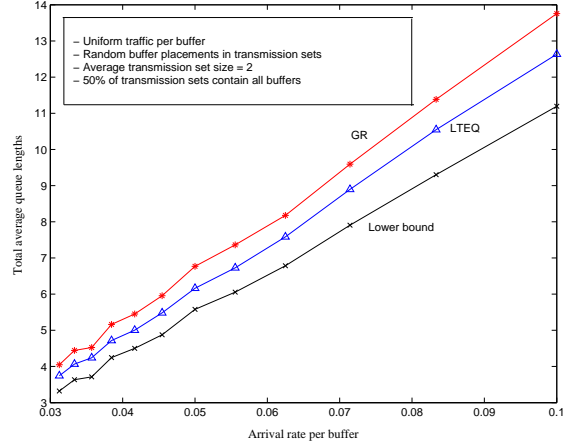
Fig. 3.   Case R1.

Fig. 4.   Case R2.

networks, Figures 3–4, the performance measures are averaged over 10 independent experiments.

One observation is that *LTEQ* outperforms *GR* in most cases. There are few cases, e.g., case D2, where *GR* is significantly better than *LTEQ* across all loads. In other cases (not presented here), dominance is alternating when load changes.

Another observation is that for some cases (D2 and R1) the lower bound is not sufficiently tight. For other cases (D1 R2), the lower bound is a very good indicator for the policy goodness. The ratio between the lower bound and the *LTEQ* performance is greater than 95% in case D1, and greater than 89% in case R2. In other cases the ratio is lower: in case D2, the ratio is around 80% and in case R1, the ratio is around 78%.

VII. CONCLUSIONS

The adaptive circuit allocation in optical circuit-switched networks has been formulated as a Markov decision process from which a lower bound to the value function has been derived. Due to problem complexity, two TDM heuristic policies, *LTEQ* and *GR*, which are well-known rules in similar queueing network contexts, are evaluated and compared with the lower bound in a variety of deterministic and random networks.

It has been observed that *LTEQ* outperforms *GR* in most cases and that the lower bound is not always tight. The best between *LTEQ* and *GR* depends on the network topology, routes and traffic. Thus, the practical conclusion of this study is that for a network in hand both policies should be evaluated using the theoretical results of Sections IV and V; then the best between

the two candidates should be selected. If the lower bound turns out to be tight, no further investigation is needed.

Note that the evaluation method derived in Section V is extremely fast compared with simulation. Furthermore, due to the vast domain of feasible policies, simulations do not present a practical option for searching the optimal policy.

REFERENCES

[1] Cisco Systems Inc., "Understanding concatenated and channelized SONET interfaces on Cisco routers," 2002. Availbale at http://www.cisco.com/warp/public/127/concat 16147.pdf

[2] Y. Chen, C. Qiao and X. Yu, "Optical Burst Switching (OBS): A New Area in Optical Networking Research," *IEEE Network Magazine*, Vol. 18 No. 3, pp. 1623, May-June 2004.

[3] W. M. Golab and R. Boutaba, "Resource Allocation in User-Controlled Circuit-Switched Optical Networks," HSNMC, LNCS 3079, pp. 776-787, Z. Mammeri and P. Lorenz (Eds.), Springer-Verlag, Berlin, 2004

[4] M. Gondran and M. Minoux, *Graphs and algorithms*, John Wiley & Sons, Inc., New Jersey, 1986.

[5] A. Itai and Z. Rosberg, "A Golden Ratio Control Policy for a Multiple-access Channel," *IEEE Trans. Auto. Contr.*, Vol. AC-29, pp. 712-718, 1984.

[6] ITU-T, Recommendation G.8080, Architecture for the automatic switched optical networks (ASON), ITU, November 2001.

[7] R. Izmailov, S. Ganguly, T. Wang, Y. Suemura, Y. Maeno, and S. Araki, "Hybrid hierarchical optical networks," *IEEE Communications Magazine*, Vol. 40, pp. 88–94, 2002.

[8] A. Jaiszczyk, "Automatically switched optical networks: benefits and requirements," *IEEE Communications Magazine*, Vol. 43, issue 2, Feb. 2005.

[9] R. M. Karp, "Reducibility among combinatorial problems," *Complexity of Computer Computations*, R. E. Miller, J. W. Thatcher Eds., Plenum Press, New York, pp. 85–103, 1972.

[10] F. P. Kelly, "Blocking Probabilities in Large Circuit-Switched Networks," *Adv. Appl. Prob.*, Vol. 18, pp. 473-505, 1986.

[11] J. M. Kleinberg, "Approximation Algorithms for Disjoint Paths Problems," Ph.D. thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, USA, 1996.

[12] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Reading, MA: Addison Wesley, 1973.

[13] A. Mokhtar and M. Azizoglu, "Adaptive Wavelength Routing in All-Optical Networks," *IEEE/ACM Trans. Networking*, Vol. 6, No. 6, pp. 197-206, Apr. 1998.

[14] B. Mukherjee, *Optical Communication Networks*, McGraw-Hill Publisher, New York, 1997.

[15] R. Ramaswami and K. N. Sivarajan, "Routing and Wavelength Assignment in All-Optical Networks," *IEEE/ACM Trans. Networking*, Vol. 3, pp. 489-500, Oct. 1995.

[16] Z. Rosberg, A. Zalesky and M. Zukerman, "Packet Delay in Optical Circuit-Switched Networks with Arbitrary Routing and Wavelength Assignment," *IEEE/ACM Transactions on Networking*, June, 2006 (in print).

[17] Z. Rosberg, A. Zalesky, H.-L. Vu and M. Zukerman, "Analysis of OBS Networks with Limited Wavelength Conversion," *IEEE/ACM Transactions on Networking*, (to appear).

[18] H.-L. Vu, Zalesky, A., Wong, E. W. M, Rosberg, Z., Bilgrami, M.S., Zukerman, M., and Tucker, R. S., "Scalable Performance Evaluation of a Hybrid Optical Switch," *Journal of Lightwave Technology*, Vol. 23, No. 10, pp. 2961–2973, Oct. 2005.