Edge-Reconfigurable Optical Networks (ERONs): Rationale, Network Design, and Evaluation

Gigi Karmous-Edwards, *Member, IEEE*, Arun Vishwanath, *Student Member, IEEE*, Douglas S. Reeves, *Member, IEEE*, Lina Battestilli, Priyanka B. Vegesna, and George N. Rouskas, *Senior Member, IEEE*

***Abstract*—To bridge the gap between the current practice of setting up expensive, dedicated, lightpath connections (i.e., static topologies), and the distant future vision of inexpensive access to dynamically switched end-to-end lightpaths, we propose a medium term solution in the form of *edge-reconfigurable optical networks (ERONs)*. An ERON is an overlay-control network created by installing readily available MEMS optical switches, and implementing a GMPLS control plane at sites interconnected by static lightpaths. The switches and control software are deployed at the edge of the network and operated by the organization-user (i.e., outside the network provider's control), hence the term "edge-reconfigurable". By providing dynamic, automated control of end-to-end lightpaths, ERONs enable the sharing of expensive network resources among multiple users and applications that require sporadic access to these resources. We develop an algorithm for creating an ERON from an existing topology of static lightpaths. We also present simulation results that quantify the benefits of ERONs, in terms of the number of lightpaths that are needed when compared to a static configuration of independent and dedicated circuits.**

***Index Terms*—Dynamic circuits, network design, optical networks.**

I. INTRODUCTION

MANY existing and emerging classes of high-end applications involve complex, intensive computations on large data sets in a manner that requires coordination of resources residing at several geographically dispersed sites. Such applications arise in a wide range of domains, including e-Science and scientific discovery, distributed simulation and visualization, petascale data mining and analytics, education, intelligence gathering, and analysis, and military applications. These applications and associated datasets are being deployed or planned on existing or emerging optical network facilities (including Internet2,[1] HOPI,[2] NLR,[3] UltraLight,[4] and TeraGrid[5]). Current research practice also tends towards collaborations among multiple groups and institutions that require high bandwidth connections to network-attached resources [5], [16]. To enable these applications, lightpaths along end-to-end paths across multi-domain networks must come up and go down, based on user requirements, and over short timescales (i.e., sub-seconds to seconds).

Ideally, the optical network would provide native support for establishing lightpaths on demand, making it possible to update the logical topology over time in response to traffic (i.e., application) demands. This model, which we refer to as the *core-reconfigurable optical network (CRON)* , encompasses the vision of a dynamically reconfigurable optical core that has been contemplated for over a decade [10]. This vision underlies the DARPA CORONET project [12] which seeks to capitalize on the maturing of optical switching technologies and the development of signaling protocol standards (e.g., GMPLS). The goal of CORONET is to prototype commercially viable approaches for network providers to offer users the ability to set up and tear down optical connections dynamically.

Nevertheless, despite years of research and development, the vision of a dynamically reconfigurable optical core is far from realization. Today, setting up high bandwidth connections on demand is nearly impossible, due to a fundamental lack of infrastructure capabilities to support the automatic and rapid establishment of end-to-end lightpaths. To establish such paths, several network providers need to be coordinated (one regional optical network at each end, as well as one or more national-scale backbone providers in the middle), and multiple contracts negotiated. Due to this administrative burden and associated timescales for the establishment of lightpaths, it is typical for lightpath connections to be held in place for long periods of time (e.g., months or longer). The result is the creation of *static topologies* consisting of a collection of independent lightpaths, each dedicated to serving a specific pair of high-end users or devices. For instance, several national research and education networks (NRENs) provide point-to-point optical connections, often reaching across more than one administrative domains, to serve existing large-scale applications [6].

Since typical applications require *sporadic* rather than continuous use of these lightpaths, the utilization of static point-to-point connections can be extremely low, often below 1% [1]. This fact is illustrated in Fig. 1 which plots the average utilization of an active 10 Gb/s interface in use within the GLIF community; similar figures are available for several such interfaces

Manuscript received December 14, 2008; revised March 30, 2009. First published April 28, 2009; current version published June 05, 2009. This work was supported in part by DARPA under Grant HR0011-08-1-0076.

The authors are with the Department of Computer Science, North Carolina State University, Raleigh, NC 27695-8206 USA (e-mail: rouskas@ncsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2009.2021279

[1][Online]. Available: http://www.internet2.edu/

[2][Online]. Available: http://networks.internet2.edu/hopi/http://networks.internet2.edu/hopi/

[3][Online]. Available: http://www.nlr.net/

[4][Online]. Available: http://www.ultralight.org/web-site/ultralight/html/index.html

[5][Online]. Available: http://www.teragrid.org/

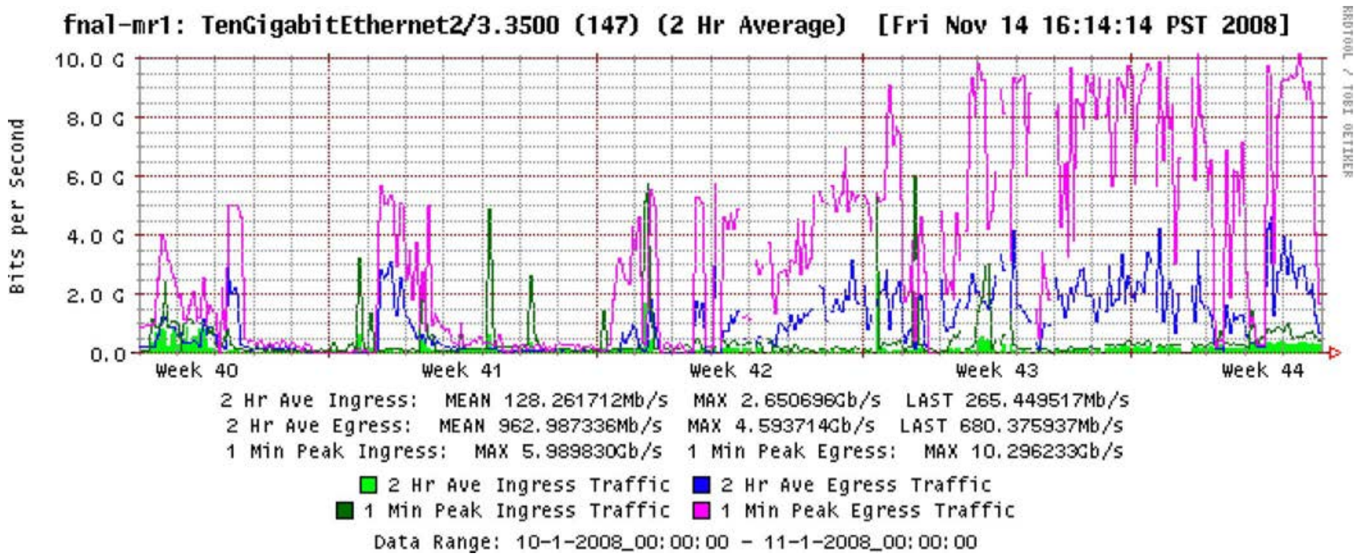

Fig. 1. Average and peak rates of a 10 Gb/s interface.

[1]. As we can see, whereas the peak transmission rate often matches the interface rate of 10 Gb/s, the average rate is significantly lower, less than 1 Gb/s for the egress and just above 100 Mbps for the ingress. These average levels roughly correspond to utilization of 73 and 7.3 hr per month, respectively. Due to the substantial expenses involved in leasing and operating long-haul lightpaths, the only viable solution is to increase utilization by sharing these resources among multiple applications and users. Lightpath sharing requires reconfigurability, i.e., the ability to use the lightpath capacity to connect dynamically a diverse pair of users, with each pair gaining exclusive access to the lightpath for a fraction of the time on an as-needed basis. To this end, several fora are hard at work to provide a solution for dynamic, inter-domain lightpath configurability, including the Open Grid Forum (OGF),[6] the Global Lambda Integrated Facility (GLIF),[7] and the Internet Engineering Task Force (IETF).[8]

To bridge the gap between the current practice of setting up expensive, dedicated, lightpath connections (i.e., static topologies), and the distant future vision of inexpensive access to dynamically switched end-to-end lightpaths, we propose a medium term solution, in the form of *edge-reconfigurable optical networks (ERONs)* . An ERON is an overlay-control network [4] created by installing readily available micro-electro-mechanical system (MEMS) optical switches and implementing a GMPLS control plane at sites interconnected by lightpaths. The switches and control software are deployed at the edge of the network and operated by the organization-user (i.e., outside the network provider's control), hence the term "edge-reconfigurable". By providing dynamic, automated control of end-to-end lightpaths, ERONs enable the sharing of expensive network resources among multiple users and applications that require occasional access to these resources. ERONs represent a practical and cost-effective solution that transforms the set of static lightpath connections owned or leased by a single organization into a flexible network topology that affords users the capability to reserve on demand, or in advance, lightpaths for any required duration.

[6][Online]. Available: www.ogf.org

[7][Online]. Available: www.glif.is

[8][Online]. Available: www.ietf.org

The remainder of this paper is organized as follows. In Section II, we explain how ERONs support dynamic lightpaths, and in Section III, we develop an algorithm for creating an ERON from an existing topology of static lightpaths. We present simulation results to quantify the benefits of ERONs in Section IV, and we conclude this paper in Section V.

II. Edge-Reconfigurable Optical Nets (ERONs)

An edge-reconfigurable optical network (ERON) consists of the following.

- A collection of permanent lightpaths that connect users at geographically dispersed sites. The lightpaths are typically leased, provide static connections among the various sites, and define a logical interconnection topology that does not change over time.
- Additional equipment at each site (i.e., at the edge of the network), including a MEMS optical switch, that operates outside the network provider's control.
- Control software, including GMPLS signaling protocols and a resource broker, that implements a control overlay among the various sites. The control software manages the MEMS switches to enable dynamic connections among users at diverse sites, allowing the sharing of the permanent lightpaths among multiple applications.

Figs. 2–4 illustrate how an organization with four sites (sites A-D in the figures, each with a number of high-end devices) might interconnect devices in different sites using a static topology, an ERON, or a CRON, respectively. In Fig. 2, the organization leases static lightpaths (denoted by solid lines), each lightpath providing connectivity between a specific pair of devices (users) at two sites. A lightpath in this case are point-to-point and dedicated to a particular pair of devices. Fig. 4 depicts the long-term CRON solution of a dynamically configurable core network. In this case, the high-end devices

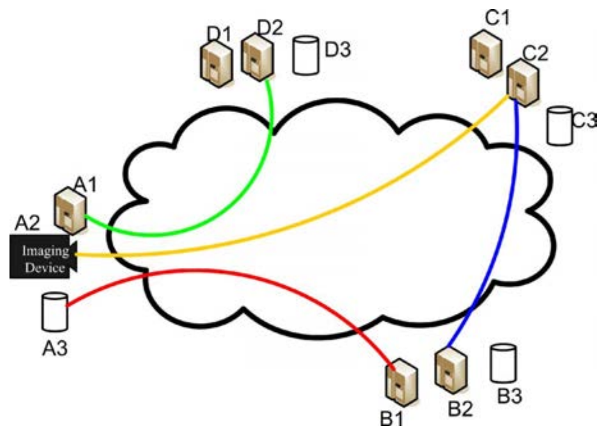

Fig. 2. Interconnecting remote sites: static topology.

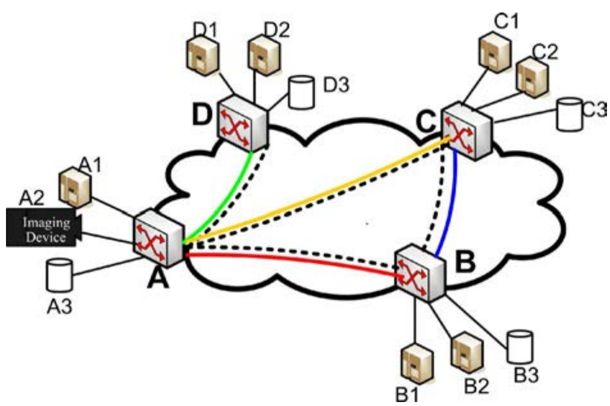

Fig. 3. Interconnecting remote sites: ERON.

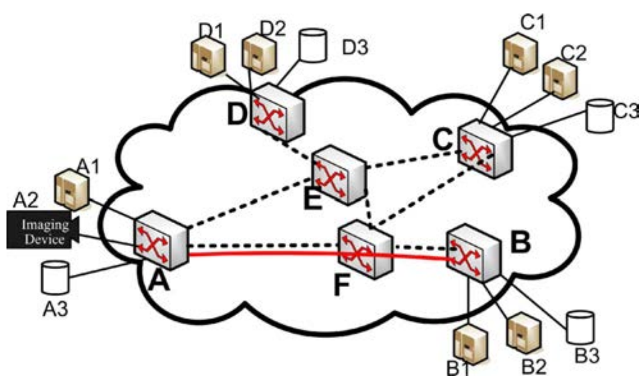

Fig. 4. Interconnecting remote sites: CRON.

at each site are attached to ingress switches that are under the network provider's control. Whenever two users at different sites wish to communicate, they request a lightpath connection from the network, which is set up dynamically and is available only for the duration of the communication. In this case, the control, setup and routing of lightpaths is the sole responsibility of the network provider; the user is not involved in this process, other than issuing lightpath requests.

Fig. 3 introduces the practical solution we refer to as ERON, which occupies the spectrum between the two extreme solutions discussed above. As in the static topology case, the organization leases from its network provider a set of static lightpaths (denoted by the solid lines in the figure). However, these lightpaths are now between the edge switches at each site, and these switches are controlled and operated by the user, *not* the provider. Under software control, the edge switches are capable of creating scheduled, dynamic connections over the static lightpaths, hence sharing the lightpath capacity among all high-end devices at the various sites.

The ERON architecture shown in Fig. 3 is inspired by the testbed of the EnLIGHTened Computing project,[9] [14], [15], a nation-wide, GMPLS-enabled testbed demonstrating the viability of the proposed approach. The underlying philosophy is to enhance the simple dedicated transport service available from today's optical network infrastructure by implementing dynamic connection functionality and associated intelligence at the edge devices. In essence, the addition of the reconfigurable edge optical switches transforms a sparse static optical network into a highly connected dynamic network.

An ERON relies on hardware and software to create and manage dynamic connections over the fixed logical topology. Each site is equipped with a fiber-based MEMS optical switch, such as the $\mathrm{DiamondWave}^{(R)}$ photonic cross-connect (PXC) product from Calient. Unlike an optical-electronic-optical (OEO) switching fabric, a MEMS optical switching fabric provides optical transparency and scalability for future network growth. Data generating devices (including compute and storage servers, cameras, telescopes, etc.) attach to the optical switch directly, or through IP routers and/or Ethernet switches, via short reach optical interfaces. The MEMS device is capable of switching any input optical signal to any of the WDM output lightpaths, and vice versa on the other end, thus creating on-demand connections among arbitrary sets of users at the various sites. Note that it is possible to set up multi-hop paths even between sites that are *not* connected directly by permanent lightpaths. Returning to Fig. 3, users at node A can be connected to users at node C by configuring the switch at node B to concatenate the lightpaths A-B and B-C to form a two-hop connection over which traffic from A to C may flow. Hence, it may not be necessary to lease expensive permanent connections between all pairs of users: by adding capacity sharing and reconfiguration capabilities, a smaller network may function as a network of much larger capacity.

The establishment and termination of the dynamic connections is performed by appropriate control software, e.g., GMPLS, at sub-second timescales. The ERON is also equipped with a resource broker, specialized software that coordinates access to the shared lightpath resources. Users contact the broker to establish or reserve lightpaths. The broker maintains an up-to-date timetable of existing reservations, and may accept, negotiate, or reject reservation requests based on resource availability, user priority, and other policy specifications.

We note that the ability to satisfy user requests is affected directly by both the amount of network resources (lightpaths) available to them and the properties of the logical topology implemented by these lightpaths. Therefore, the issue of logical topology design is integral to the design and implementation of ERONs, and is the subject of the next section.

[9][Online]. Available: www.enlightenedcomputing.org

III. ERON Topology Design

The problem of designing logical topologies of lightpaths given some information about the traffic demands among nodes of an optical network has been studied extensively in the literature; the interested reader is referred to [2], [3], [7]–[9], [11], [13], [17], and references thereof. The ERON topology design problem we consider in this paper differs from previous studies in several respects. First, ERONs are overlay networks, whereas earlier work dealt with the design of core networks. Second, most algorithms were developed for constructing a logical topology from scratch, whereas our approach is to build an ERON starting from an existing static topology. More importantly, the objective of most logical topology studies has been to minimize the maximum congestion, which is achieved by balancing the traffic load over the network links. On the other hand, our objective is to minimize blocking probability, which requires that the load be concentrated on a small number of high-capacity links. The more recent work in [18] is similar to ours in that it compares static lightpath topologies to dynamic optical networks. Although some of the conclusions of [18] are consistent with ours, our network and traffic models and design approach are more representative of realistic scenarios.

A. Current Practice: Static Topology of Lightpaths

Consider an organization with M users (i.e., high-end devices) distributed across $N < M$ geographic locations. In current practice, if two users at different locations wish to communicate, a permanent *end-to-end (e2e)* lightpath must be established between them. In the absence of switching capabilities, the e2e lightpath is *dedicated* to communication between these users only, and cannot be accessed by other users at the same locations. We assume that e2e lightpaths are unidirectional, hence, two e2e lightpaths, one in each direction, must be set up if bidirectional communication is required.

We make a distinction between e2e and *intra-domain* lightpaths. Since users of a national or global organization may reside at locations in different administrative domains, an e2e lightpath between two users may consist of a string of intra-domain lightpaths, each originating and terminating at the boundaries of a single administrative domain. We assume that the organization leases intra-domain lightpaths separately from each network provider, and connects (patches) them together at its own premises (i.e., at the edge, outside the operator's network) to form longer e2e lightpaths whenever necessary.

In the absence of reconfigurable elements inside the network, providing full connectivity among all M users would require $M(M-1)$ permanent e2e lightpaths, each dedicated to a directed pair of users. Since leasing such a number of e2e lightpaths would be impractical and prohibitively expensive for other than a small number M of users, the organization might lease a smaller number K of e2e lightpaths between a select set of user pairs; typically, $K \ll M(M-1)$. The K pairs of users to be connected directly may be selected based on the amount or criticality of information exchanged, the priority of users, or some other relevant criteria. Let t_k denote the long-term traffic demand for the kth user pair, expressed in gigabits per second. Let C denote the lightpath capacity, also expressed in gigabits per second. Then, the kth pair is assigned $\lceil t_k/C \rceil$ e2e lightpaths.

Algorithm for Removing Capacity from a Static Topology
Input: Static topology of L intra-domain links
Output: An ERON topology with fewer lightpaths

```
begin
c_l ← capacity of link l, l = 1, ..., L, in lightpaths
Run simulation, obtain link utilizations {u_l} and BP b
while b < 10^-3 do
    for l = 1 to L do // Consider each link in isolation
        ρ_l ← Erl^-1(10^-4, c_l) // Target link offered load
        u'_l ← ρ_i / c_l // Corresponding target link utilization
        slack_l ← u_l / u'_l
    endfor
    j ← link such that slack_j = min_l{slack_l}
    c_j ← c_j − 1      // Reduce capacity of link j
    Run simulation, obtain link utilizations {u_l} and BP b
endwhile
return the last topology for which b < 10^-3
end
```

Fig. 5. Algorithm for constructing an ERON topology by removing capacity from the links of the initial static topology.

Note that an e2e lightpath may consist of multiple intra-domain lightpaths, and all e2e lightpaths routed over a given intra-domain hop between some sites A and B require a dedicated intra-domain lightpath from A to B. Let L be the number of distinct intra-domain hops required by the e2e lightpaths and c_l be the number of intra-domain lightpaths set up over the l-th hop. These L hops are the links of the *static topology* defined among the K pairs of users. The capacities of the L links are such that $\sum_{l=1}^{L} c_l = \sum_{k=1}^{K} h_k \lceil t_k/C \rceil$, where h_k is the number of intra-domain lightpaths making up the e2e lightpath between the kth user pair.

B. ERON Topology Algorithm

Dedicated lightpaths are expensive resources that may be economically justified only if their utilization remains at high levels throughout the lease period. Recent statistics (refer to Fig. 1) indicate that the utilization of GLIF lightpaths is quite low, often as low as a few hours per month [1]. Alternatively, an organization may deploy an ERON to increase the utilization of its leased lightpaths. This increase is due to two factors: 1) higher connectivity, as all users at a given site have access to all incoming and outgoing lightpaths and 2) sharing of the (previously dedicated) lightpaths among this larger number of user pairs.

An important question that arises is whether deploying an ERON would result in significant savings in the number of lightpaths to be leased to justify the (one-time) hardware and software expense. Note that, by allowing pairs of users to share a lightpath, an ERON introduces the possibility of blocking. To ensure that users receive a good quality of service, we assume an acceptable upper bound of 10^{-3} in *network-wide* blocking probability.

In order to quantify the benefits of ERON in terms of the number of lightpaths required to provide connectivity among the same K pairs of users connected by the static topology, we consider the following problem.

1) Problem 3.1: Given a static topology of L intra-domain links connecting K pairs of users, the (unprotected) long-term traffic demands $t_k, k = 1, \ldots, K$, and the capacity C of a

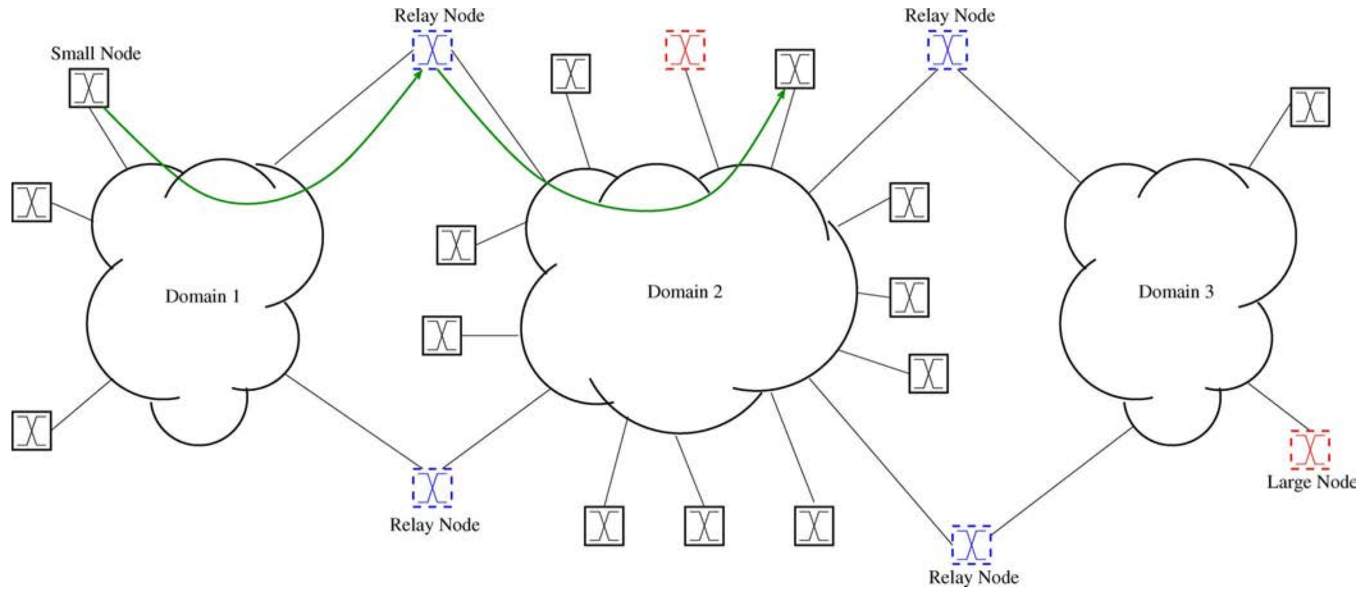

Fig. 6. Three-domain ERON network.

lightpath, design an ERON topology with the least number of lightpaths such that the network-wide blocking probability (BP) $b \leq 10^{-3}$.

Logical topology design problems are typically NP-hard [3]. An additional challenge in this case is the difficulty of expressing the blocking probability in exact and closed-form manner. We also emphasize that the goal of this study is not to identify optimal solutions but rather to quantify the benefits that are practically achievable using the ERON model. Hence, to solve the above problem we use the greedy algorithm presented in Fig. 5. The algorithm starts with the static topology and selectively removes lightpaths as long as the blocking probability remains below 10^{-3}.

The algorithm first runs a simulation of the ERON network with a topology identical to the one of the static topology (Step 2 in Fig. 5). This simulation attempts to route user demands over a candidate network design, in order of arrival of the demands; more details are provided in Section IV-C. From the simulation, we obtain the overall blocking probability in the ERON, as well as the utilization u_i of each link i in the network. The main operation of the algorithm consists of the **while** loop between Steps 3 and 12, which is executed as long as the overall blocking probability is below the threshold of 10^{-3}. Considering each link i of the network in isolation, Step 5 determines a target offered load ρ_i to the link that would result in the blocking probability on this link (computed using the Erlang-B formula with the number of servers equal to the capacity c_i of the link) being equal to 10^{-4}[10] The corresponding target link utilization u'_i is obtained by dividing the target offered load by the capacity c_i of the link (Step 6). The *slack* of a link is then defined as the ratio of its actual utilization u_i (obtained through simulation) and the target utilization u'_i, as shown in Step 7. Note that a low slack value implies excess capacity at a link relative to the actual (simulation) offered load. Therefore, in Steps 9 and 10, the algorithm reduces the capacity of the link with the minimum slack by one lightpath. The algorithm then runs the simulation again, and repeats the **while** loop to reduce capacity further, if the blocking probability remains below 10^{-3}.

[10]We have found that this target link probability is sufficient to guarantee the network-wide blocking probability for networks with a small diameter.

IV. Performance Evaluation

A. Network Model

We consider the three-domain ERON network shown in Fig. 6. The network consists of $N = 20$ ERON switches (nodes), each representing a site of a global organization. Attached to the ERON nodes are M *users* that generate the traffic demands; note that the users are *not* shown in Fig. 6. The users are distributed across the various ERON nodes based on the size of each node, as we describe shortly.

The ERON nodes belong to one of three domains based on their geographic locations. For instance, the three domains in Fig. 6 might correspond roughly to the providers serving an organization's sites in Asia, US, and Europe, respectively. The middle domain is the largest and contains about 70% of the ERON nodes. The side domains are smaller, with each containing about 15% of the nodes. Lightpaths terminate at domain boundaries. Thus, an e2e lightpath that connects a node in Domain 1 to a node in Domain 3 is made up of three intra-domain lightpaths. For the topology in Fig. 6, an e2e lightpath may consist of one, two, or three intra-domain lightpaths, depending on the location of its endpoints.

ERON nodes are classified as either *small* or *large* . Large nodes, denoted by dotted lines in Fig. 6, have more users connected to them than small nodes. Consequently, the total aggregate traffic demand from/to large nodes is higher than from/to small nodes. We assume that 6 of the ERON nodes are large and 14 are small. The three domains are interconnected by four of the large ERON nodes, referred to as the *relay* nodes. Just as any other large node, relay nodes have users connected to them that generate traffic. Unlike other large nodes, however, each relay node is attached to two adjacent domains. This requires that each is equipped with an OEO capability, and the relay node terminates all intra-domain lightpaths that are part of an e2e lightpath passing through it. This capability is necessary so that the organization may concatenate lightpaths provided by independent service providers to form longer e2e lightpaths as needed, without intervention from the service providers.

B. Traffic Model

We assume that a lightpath is unidirectional and its capacity is $C = 10$ Gb/s. User demands that exceed 10 Gb/s will be assigned multiple lightpaths when they are routed. Let D_{tot} denote the aggregate network traffic, in gigabits per second, generated by users attached to ERON nodes. We denote the aggregate traffic generated by large-to-large, large-to-small, and small-to-small connections as D_{ll}, D_{ls} and D_{ss}, respectively; clearly, $D_{\mathrm{tot}} = D_{ll} + D_{ls} + D_{ss}$. Let K be the number of (unidirectional) connections that make up the static topology from which the ERON topology is derived. We assign each of the $2K$ endpoints (users) of these connections probabilistically, such that users are twice as likely to be assigned to a large node as to a small node. The total amount of traffic D_{tot} is then divided into components D_{ll}, D_{ls} and D_{ss} in proportion to the number of connections (out of K) that fall into the large-to-large, large-to-small, and small-to-small classes, respectively.

We also consider three traffic patterns: *uniform*, *distance-decreasing*, and *distance-increasing*. The individual traffic components are then distributed over the connections within each class according to one of these traffic patterns: in the uniform pattern, the traffic is divided equally among all connections in the same class, while in the distance-decreasing (respectively, distance-increasing) case, the traffic is divided among connections in a class in inverse proportion (respectively, direct proportion) to the distance between the source and destination nodes. The result is a list of K user connections and their specific (average) data rates.

We consider two scenarios. In the *low* traffic scenario, the average aggregate traffic D_{tot}[11] over all connections does not exceed 300 Gb/s, whereas in the *high* traffic scenario, the average aggregate traffic varies between 1 and 3 Tb/s.

C. Simulation Design and Methodology

To carry out the evaluation study, we implemented our own event-driven simulator in C. Connection requests were generated according to a Poisson arrival process. Each traffic demand has a holding time (duration) and a traffic amount. Each traffic demand amount was in units of lightpaths, ranging from a low of 1 lightpath to a high of 4 lightpaths. Traffic amounts were generated according to a truncated power-law distribution, using an exponent of 1.5. The holding times of demands normally ranged from 10 s to 1 hr. These demands were probabilistically generated according to a truncated Pareto distribution, again with an exponent of 1.5.

[11]We distinguish between *average* and *peak* traffic on a connection. The peak traffic may fully utilizing the capacity of a lightpath, whereas the average traffic can be quite low, especially if the specific connection is used sporadically, e.g., for a few hours per month. The ERON topology algorithm we described earlier is based on average traffic demands, not peak ones.

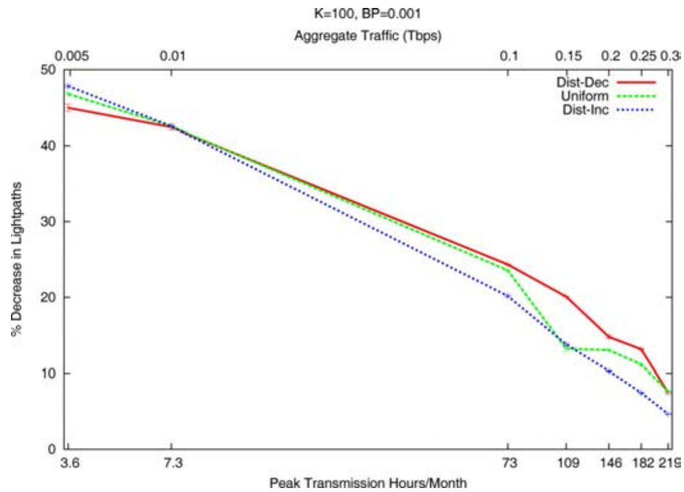

Fig. 7. Low traffic: effect of average aggregate traffic amount D_{tot} (and equivalent peak transmission hours/month), $K = 100$.

We generated and simulated the routing of traffic that conforms to the average rates for each pair of users dictated by the traffic model above. This was accomplished by setting the value of the rate λ of the Poisson arrival process for each pair of users to a value such that the aggregated average matches the corresponding element in the traffic matrix.

Each simulation run lasted until a set of 500 000 traffic demands was generated and routed. For every data point shown on the graphs, 10 runs were executed, each with a different random seed. From these, the mean and the 95% confidence intervals were computed (shown on the graphs as error bars).

The routing algorithm simulated is a standard method of achieving good overall network utilization and low blocking. First, from among the available paths to the destination, only paths that had sufficient available (i.e., residual, not currently used) capacity were considered. From among such paths, the shortest (in terms of hop count) path was chosen. In the event of a tie, one path from among the shortest, highest available capacity ones was selected uniformly randomly.

D. Numerical Results

We present a set of experiments to quantify the benefits of ERONs compared to establishing a static topology of dedicated connections among a given set of users. Fig. 7 shows the achievable decrease in number of intra-domain lightpaths for the case of $K = 100$ connections. Each data point in the figure is an average over 10 problem instances generated according to one of the three traffic patterns we described earlier; 95% confidence intervals are also plotted. In this case, the static topology always consists of $K = 100$ e2e lightpaths (each made up of at least one intra-domain lightpath) between a set of users; this set is randomly generated for each problem instance and each user is probabilistically assigned to one of the ERON nodes, as we explained above. For a given instance, the ERON topology is constructed using the algorithm in Fig. 5 to remove capacity from the static topology.

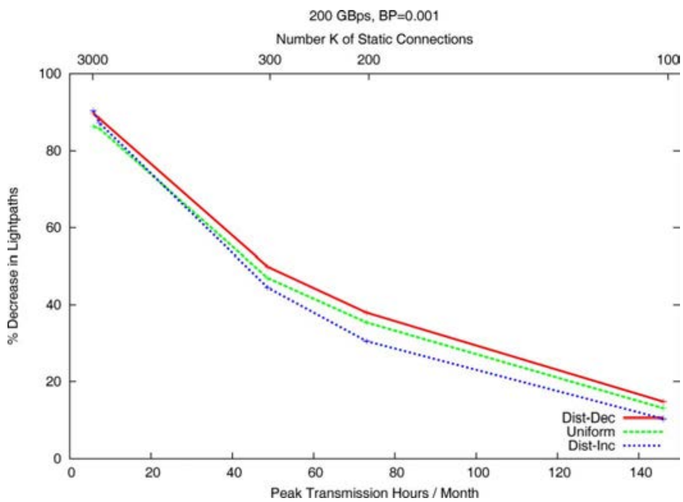

Fig. 8. Low traffic: effect of number K of connections (and equivalent peak transmission hours per month), $D_{\rm tot} = 200$ Gb/s.

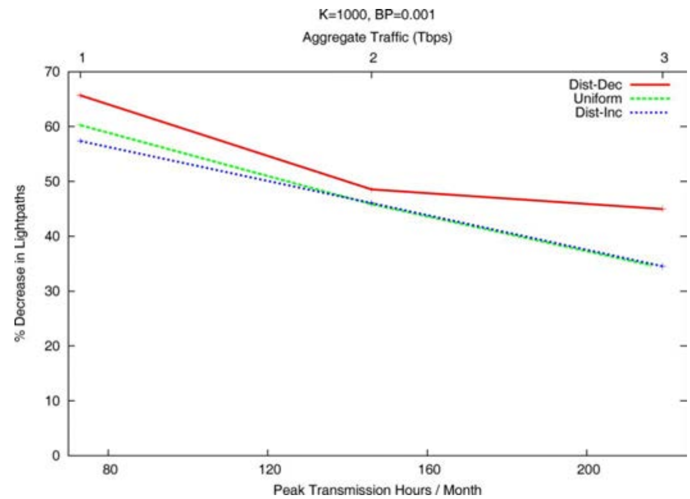

Fig. 9. High traffic scenario: effect of average aggregate traffic amount $D_{\rm tot}$ (and equivalent peak transmission hours per month), $K = 1000$.

The figure plots the decrease in lightpaths against the average aggregate traffic $D_{\rm tot}$ generated by the K connections. To put the aggregate traffic values in perspective, note that the $K = 100$ static connections can generate a peak traffic equal to 1 Tb/s (since the capacity of each e2e lightpath is 10 Gb/s). Consequently, an average aggregate traffic $D_{\rm tot} = 0.01$ Tb/s corresponds to each connection transmitting at its peak rate for an average of 7.3 hr per month (and being idle the remaining of the time). At low utilization, the savings can be substantial, more than 40% compared to the static topology. As utilization increases, the savings decrease accordingly, but even when $D_{\rm tot} = 0.1$ Tb/s, the savings can exceed 20%. Also, the traffic pattern does affect the results, but not substantially.

For the results shown in Fig. 8, we let the average aggregate traffic $D_{\rm tot} = 200$ Gb/s, and vary the number K of connections in the static topology. In this case, $K = 2000$ corresponds to each connection transmitting at peak rate for an average of 7.3 hr per month. Again, we observe that at utilization levels of 7.3 hr per month or lower, the achievable savings in lightpaths are quite significant, between 80%–90% compared to the static topology. Even at higher utilization levels, e.g., $K = 200$, the savings are about 30% or more.

Figs. 9 and 10 are similar to Figs. 7 and 8, respectively, but present results for the high traffic scenario. In Fig. 9 we let $K = 1000$ connections and vary the average aggregate amount $D_{\rm tot}$ of traffic from 1–3 Tb/s. Note that $D_{\rm tot} = 1$ Tb/s corresponds to each connection transmitting at peak rate of 10 Gb/s for an average of 73 hours per month, and remaining idle the rest of the time. At this level of utilization, the average savings in terms of lightpaths is approximately between 55%–65%, depending on the traffic pattern, decreasing to between 35%–45% as utilization increases to 219 hours per month (for $D_{\rm tot} = 3$ Tb/s). In Fig. 10 we let $D_{\rm tot} = 2$ Tb/s, and vary the number K of connections from 1000 (corresponding to utilization of 146 hours per month) to 3000 (utilization of about 49 hr per month). Again, as utilization decreases, the average lightpath savings increase substantially, reaching 75% or more over the static topology for $K = 3000$.

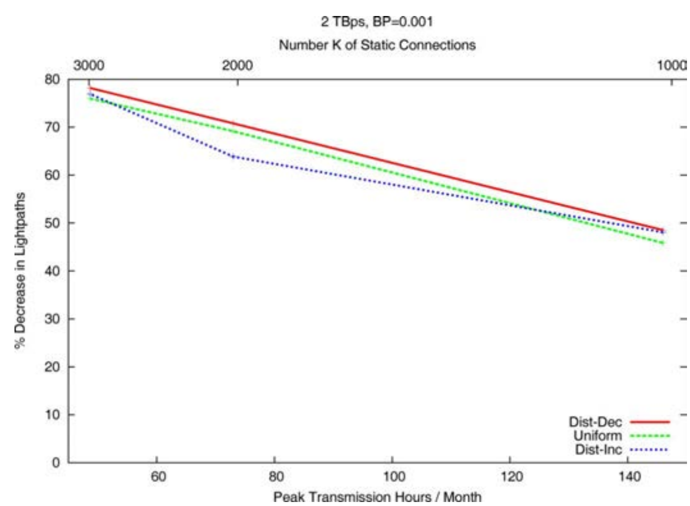

Fig. 10. High traffic scenario: effect of number K of connections (and equivalent peak transmission hours/month), $D_{\rm tot} = 2$ Tb/s.

By comparing Figs. 9 and 10 (high traffic scenario) to Figs. 7 and 8, respectively (low traffic scenario), we observe that under the high traffic scenario it is possible to achieve substantially more savings at the same level of utilization (equivalently, achieve the same amount of savings at higher utilization). Consider, for instance, the savings for $D_{\rm tot} = 0.1$ Tb/s in Fig. 7 and for $D_{\rm tot} = 1$ Tb/s in Fig. 9 and for $D_{\rm tot} = 1$ Tb/s; both correspond to a utilization level of 73 hr per month. The savings in Fig. 7 are between 20%–25%, whereas in Fig. 9 are significantly higher, between 45%–50%, depending on the traffic pattern. This result is due to the fact that in both cases the capacity of a lightpath is the same (i.e., 10 Gb/s), but under the high traffic scenario, the links of the ERON topology have much higher capacity (number of lightpaths) than under the low traffic scenario. As a result, the blocking probability of the larger network can be substantially lower under the same offered load.

V. Concluding Remarks

We have proposed ERONs, overlay dynamic optical networks based on a collection of static lightpaths and MEMS optical switches with GMPLS control. An ERON enables the sharing of expensive lightpaths by providing reconfigurability at the edges of the optical network. We consider ERONs as a medium-term solution for enabling today's emerging high-end applications in the areas of science, defense, and enterprise. Simulation results demonstrate that, depending on average utilization, migration to an ERON network from a static topology may constitute a significant savings in the number of static lightpaths required to meet the needs of the traffic demands at an acceptable blocking probability.

Acknowledgment

The authors would like to thank A. Saleh, J. M. Simmons, and R. Henry for their invaluable insights and suggestions.

References

[1] C. Guok, "Private Correspondence with Sr Network Engineer," ESnet, Lawrence Berkeley National Laboratory, 2008.

[2] J. A. Bannister, L. Fratta, and M. Gerla, "Topological design of the wavelength-division optical network," in *Proc. INFOCOM*, San Francisco, CA, 1990, pp. 1005–1013.

[3] R. Dutta and G. N. Rouskas, "A survey of virtual topology design algorithms for wavelength routed optical networks," *Opt. Netw.*, vol. 1, no. 1, pp. 73–89, Jan. 2000.

[4] A. Jukan and G. Karmous-Edwards, "Optical control plane for the grid community: A tutorial," *IEEE Commun. Surveys Tutorials*, vol. 9, no. 3, pp. 30–44, 2007.

[5] G. Karmous-Edwards, "Global E-Science collaboration," *IEEE Comput. Scientific Eng.*, vol. 7, no. 2, pp. 20–27, Mar./Apr. 2004.

[6] G. Karmous-Edwards, "Today's optical network research infrastructure for E-Science applications," presented at the OFC, Anaheim, CA, 2006.

[7] R. Krishnaswamy and K. Sivarajan, "Design of logical topologies: A linear formulation for wavelength routed optical networks with no wavelength changers," in *Proc. INFOCOM*, 1998, pp. 919–927.

[8] E. Leonardi, M. Mellia, and M. A. Marsan, "Algorithms for the logical topology design in WDM all-optical networks," *Opt. Netw.*, vol. 1, no. 1, pp. 35–46, Jan. 2000.

[9] R. Ramaswami and K. N. Sivarajan, "Design of logical topologies for wavelength-routed optical networks," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 840–851, Jun. 1996.

[10] A. A. Saleh and J. M. Simmons, "Evolution toward the next-generation core optical network," *IEEE/OSA J. Lightw. Technol.*, vol. 24, no. 9, pp. 3303–3321, Sep. 2006.

[11] S. Sengupta and R. Ramamurthy, "From network design to dynamic provisioning and restoration in optical cross-connect mesh networks: An architectural and algorithmic overview," *IEEE Network*, vol. 15, no. 4, pp. 46–54, Jul./Aug. 2001.

[12] Defense Advanced Research Projects Agency (DARPA), Location: Contracts Management Office, Washington, DC, "Dynamic Multi-Terabit Core Optical Networks: Architecture, Protocols, Control and Management (CORONET)," 2006. [Online]. Available: http://www.darpa.mil/STO/Solicitations/CORONET/index.htm

[13] B. Mukherjee, D. Banerjee, S. Ramamurthy, and A. Mukherjee, "Some principles for designing a wide-area WDM optical network," *IEEE/ACM Trans. Netw.*, vol. 4, pp. 684–696, Oct. 1996.

[14] L. Battestilli, , A. Hutanu, G. Karmous-Edwards, D. S. Katz, J. MacLaren, J. Mambretti, J. H. Moore, S.-J. Park, H. G. Perros, S. Sundar, S. Tanwir, S. R. Thorpe, and Y. Xin, "EnLIGHTened computing: An architecture for co-allocating network, compute, and other grid resources for high-end applications," presented at the IEEE Honet, Dubai, UAE, Nov. 2007.

[15] S. Thorpe *et al.*, "G-lambda and EnLIGHTened: Wrapped in middleware coallocating compute and network resources across Japan and the US," presented at the Proc. Gridnets, Lyon, France, Oct. 2007.

[16] *Grid Networks: Enabling Grids with Advanced Communication Technology*, F. Travostino, J. Mambretti, and G. Karmous-Edwards, Eds. New York: Wiley, 2006.

[17] N. Wauters and P. Demeester, "Design of the optical path layer in multiwavelength cross-connected network," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 881–892, Jun. 1996.

[18] A. Zapata and P. Bayvel, "Dynamic vs. static wavelength-routed optical networks," *J. Lightw. Technol.*, to be published.

Gigi Karmous-Edwards received the B.S. degree in chemical engineering and the M.S. degree in electrical engineering, from North Carolina State University (NCSU) and is currently enrolled in a Ph.D. program.

She is a Principal Scientist at MCNC and Adjunct Professor at NCSU, Raleigh. She has been active in research projects on the control and management of optical networks in Grid environments. She has published numerous papers and coauthored the book *Grid Networks* (Wiley, 2006). She served as PI of the NSF seed-funded large-scale international collaborative project, Enlightened Computing, on virtualization of Grid resources for E-science applications). She was also a co-PI for the DARPA project entitled ERON. She is co-chair of the Technical and Control plane working group of the Global Lambda Infrastructure Facility (GLIF). She served on technical advisory board of EU's EARNEST for Geant2 and LONI. She was chair of GridNets 2005, served on the organizing committee of Broadnets 2007, 2008, and she organized and chaired two International workshops for Optical Control Plane for the Grid Community, in 2004.

Arun Viswanath is pursuing the Ph.D. degree in the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia.

He was a visiting Ph.D. student in the Department of Computer Science at North Carolina State University in 2008. His research interests include the area of optical networks, QoS routing, and router buffer sizing.

Douglas S. Reeves received the Ph.D. degree in computer science from the Pennsylvania State University, University Park.

He is a Professor of Computer Science with North Carolina State University, Raleigh. His research interests include networks, distributed computing, and security.

Tzvetelina Lina Battestilli received the B.S. degree in electrical engineering from Kettering University, Flint, MI, in 1999, the M.S. degree in computer networking and the Ph.D. degree in computer science from North Carolina State University (NCSU), Raleigh, in 2002 and 2005, respectively.

She is a researcher in the Next Generation Computing Systems Group, IBM, Research Triangle Park. NC. Her research has been in the field of computer networking with special emphasis on network design, dynamic optical networks, and Grids.

Priyanka B. Vegesna received the Masters degree in computer networking from the Electrical and Computer Engineering Department, North Carolina State University, Raleigh, in 2008.

Her research interests include network simulation, optical networks, and wireless networks.

George N. Rouskas (S'92–M'95–SM'01) received the Diploma in computer engineering from the National Technical University of Athens (NTUA), Athens, Greece, in 1989, and the M.S. and Ph.D. degrees in computer science from the College of Computing, Georgia Institute of Technology, Atlanta, GA, in 1991 and 1994, respectively.

He is a Professor of Computer Science at North Carolina State University, Raleigh. During the 2000–2001 academic year he spent a sabbatical term at Vitesse Semiconductor, Morrisville, NC, and in May 2000, December 2002, and July 2006 he was an Invited Professor at the University of Evry, France. His research interests include network architectures and protocols, optical networks, multicast communication, and performance evaluation. He is author of the book *Internet Tiered Services: Theory, Economics, and Quality of Service* (Springer, 2009) and coeditor of the book *Traffic Grooming for Optical Networks: Foundations, Techniques and Frontiers* (Springer, 2008).

Dr. Rouskas was a recipient of the 2004 ALCOA Foundation Engineering Research Achievement Award, the 2003 NCSU Alumni Outstanding Research Award, a 1997 NSF Faculty Early Career Development (CAREER) Award, and the 1994 Graduate Research Assistant Award from the College of Computing, Georgia Tech. He is a coauthor of two papers that received Best Paper Awards at the 1998 SPIE conference on All-Optical Networking and at the 2006 CSNDSP conference. He is especially proud of his teaching awards, including his induction in the NCSU Academy of Outstanding Teachers in 2004, and the Outstanding New Teacher Award he received from the Department of Computer Science in 1995. He is founding Coeditor-In-Chief of *Optical Switching and Networking* (OSN), an Elsevier journal, he has served on the editorial boards of the IEEE/ACM TRANSACTIONS ON NETWORKING, *Computer Networks*, and *Optical Networks*, and he was a coguest editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He was general co-chair of IEEE BROADNETS 2007, technical program co-chair of the Networking 2004 conference, general co-chair of the IEEE LANMAN 2005 workshop, program chair of the IEEE LANMAN 2004 workshop, and program co-chair of the Traffic Grooming workshop 2004. He is a member of the ACM and of the Technical Chamber of Greece.