INTRODUCTION

The term **optical networking** has evolved to embody a broad range of technologies over the last two decades [1]. From the late 1980s to early 1990s, this terminology was synonymous with legacy time division multiplexing (TDM), synchronous optical network/synchronous digital hierarchy (SONET/SDH), and rigid transmission formats/mappings. The Internet boom of the mid-1990s saw this designation shift to dense wavelength division multiplexing (DWDM). This paradigm represented a fundamental advance as it opened up unused single-mode fiber (SMF) C/L band spectrum. Nevertheless, large-granularity DWDM pipes never replaced TDM in smaller metro/access markets, which instead gravitated toward agile next-generation SONET/SDH (NGS) sub-rate grooming. In all, a rich set of dynamic optical capabilities have now materialized, enabled by new optical cross-connect (OXC), reconfigurable add-drop multiplexer (ROADM), and multiservice provisioning (MSP) designs. Furthermore, hybrid systems are packaging multiples of such functionalities into a single node.

In light of the above, optical networks are evolving into a complex interconnection of circuit-switched domains and layers (or granularities), with the delineations being driven by many factors, for example, geographic, administrative, client requirements, economic cost, entrenched infrastructure, and so on [2]. Concurrently, the scale and reach of high-bandwidth applications continues to grow, mandating services delivery across heterogeneous optical domains. A fine example is the e-science community, which already uses optical networks to support multigigabit connectivity for applications such as remote steering, high energy physics, climate modeling, genetics, and so on. Commercial Ethernet private line (EPL) and virtual local area network (VLAN) services also pose similar demands. In all, these trends are driving the need for multidomain, multilayer optical control plane integration, termed *vertical-horizontal integration*; see Fig. 1.

From a standards perspective, much progress has been made in optical control via frameworks such as the Internet Engineering Task Force (IETF) generalized multiprotocol label switching (GMPLS), the International Telecommunication Union-Telecommunication (ITU-T) automatic switched transport network (ASTN), and the Optical Internet Form (OIF) user network interface (UNI) and network-network interface (NNI) [1]. Although these standards provide varying degrees of multidomain support, related algorithm design/performance evaluations are largely missing. Instead, most work has treated single-domain renditions of ubiquitous problems such as routing and wavelength assignment (RWA), virtual topology design, survivability, grooming, and so on. Clearly, the extension of these schemes to distributed multidomain (multicarrier) settings poses many concerns. Foremost, it is not feasible for a single entity to maintain a global state due to obvious scalability and confidentiality concerns. Although proprietary solutions are being developed to integrate multivendor operational support systems (OSS), these have limited scope [2]. Therefore, interdomain interconnection continues to be performed manually, yielding high inefficiency and long lead times.

ABSTRACT

As optical networks proliferate, there is a growing need to address distributed multidomain provisioning. Although multi-domain operation has been well-studied in packet/cell-switching networks, the multilayer (granularity) circuit-switched nature of modern optical networks presents a unique set of challenges. This survey addresses control plane design for such heterogeneous infrastructures and describes new challenges in the areas of state dissemination, path computation, and survivability. Sample results from a recent study also are presented.
Note that asynchronous transfer mode (ATM) represents multiple "abstract" nodes and links. At higher levels, an RA represents a domain comprised of routing areas (RAs). At the lowest hierarchical level, an RA represents a domain comprised of physical nodes and links. At higher levels, an RA represents multiple "abstract" nodes and links. Note that asynchronous transfer mode (ATM) also defined a hierarchical design with peer groups, namely, private network-to-network interface (PNNI) protocol. However, ASTN further defines component groups to set up, maintain, and release connections; for example, an RA can have one or more routing controller (RC) entities. Associated component functions also are outlined for tasks such as auto-discovery, auto-provisioning, restoration, and so on. In ASTN, network topology is not made visible to the client layer, and hence, connections are treated as subnetwork point pool (SNPP) links. Overall, ASTN is quite flexible because each lower-layer control plane can be tailored to the particular type of equipment (layer). Nevertheless, because ASTN defines only architectures, its liaison efforts with the IETF and OIF are of crucial importance.

**Figure 1.** Multidomain/multilayer optical networks.

In light of the above, comprehensive multidomain/multilayer optical standards and algorithms are required for a distributed operation with limited global visibility, namely, resource dissemination, traffic engineering (TE) path computation, and survivability. These solutions must be scalable and achieve a level of optimality, where optimality can be inferred as the TE path chosen in the idealized case of a “flat” network, that is, no partitioning with a global state [3]. Typically, these objectives can be conflicting. Currently, although multidomain data networks were well studied, optical networks pose very different resource/link constraints and have an inherent grooming aspect due to multiple granularities. This article addresses this challenging area and is organized as follows. We first present a survey of existing standards. Next, we describe related research work and highlight key open challenges. Then, a sample multidomain DWDM study is presented along with conclusions.

**MULTIDOMAIN OPTICAL NETWORKING STANDARDS**

Various optical standards were developed within the IETF, ITU-T, and OIF. In this section, we survey these efforts, with a particular focus on their multidomain capabilities; see also Table 1.

**ITU-T**

The ITU-T has been maturing its multidomain-capable ASTN framework for several years (G.8080, formerly G.asn) [1]. The reference architecture here defines a hierarchical set up of routing areas (RAs). At the lowest hierarchical level, an RA represents a domain comprised of physical nodes and links. At higher levels, an RA represents multiple “abstract” nodes and links. Note that asynchronous transfer mode (ATM) is quite flexible because each lower-layer control plane can be tailored to the particular type of equipment (layer). Nevertheless, because ASTN defines only architectures, its liaison efforts with the IETF and OIF are of crucial importance.
Table 1. Multidomain optical networking standards.

<table>
<thead>
<tr>
<th>Body</th>
<th>Framework</th>
<th>Routing</th>
<th>Signaling</th>
<th>TE path comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITU-T</td>
<td>Automatically switched transport network (ASTN), formerly ASON (G.8080)</td>
<td>• G.7713.1/Y.1704: distributed call and connection management (PNNI-based)</td>
<td>• G.7713.2/Y.1704: distributed call and connection management (GMPLS RSVP-TE-based)</td>
<td>• None specified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• G.7714/Y.1705: generalized automatic discovery</td>
<td>• G.7713.3/Y.1704: distributed call and connection management (GMPLS CR-LDP-based)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• G.7715/Y.1706: architecture and requirements for routing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IETF</td>
<td>Generalized multiprotocol label switching (GMPLS)</td>
<td>• RFC 4258 (requirements for GMPLS for ASON)</td>
<td>• RFC 4208 GMPLS UNI</td>
<td>• Policy-enabled PCE framework</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• IGP routing extensions for discovery of TE node capabilities</td>
<td>• RSVP-TE for overlay model</td>
<td>• PCE protocol (PCEP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• OSPF extensions in support of inter-AS MPLS and GMPLS</td>
<td>• Interdomain MPLS and GMPLS TE extensions for RSVP-TE</td>
<td>• Per-domain path computation for inter-domain TE LSP setup</td>
</tr>
<tr>
<td>OIF</td>
<td>User-network interface (UNI), network-network interface (NNI)</td>
<td>• OIF UNI 2.0</td>
<td>• OIF-E-NNI-Sig-2.0: intracarrier E-NNI signaling</td>
<td>• None specified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• E-NNI OSPF-01.0</td>
<td>• OIF-UNI-01.0-R2-RSVP: RSVP extensions for UNI 1.0</td>
<td></td>
</tr>
</tbody>
</table>

erability. Recently, the OIF also detailed routing and signaling functionalities for E-NNI. Specifically, a hierarchical routing setup is defined (ASTN G.8080), based upon open shortest path first traffic engineering (OSPF-TE). However, the inter-carrier case has not been fully addressed yet. Overall, UNI and NNI can automate circuit setups across multiple optical layers, both DWDM and SONET/SDH. Note that the use of link state routing in multidomain optical settings is quite germane as the number of domains will be drastically lower than the number of autonomous systems (ASs) in the wider Internet (which uses distance-vector routing).

**IETF**

Internet Protocol (IP) networks feature a mature multidomain setup comprised of a hierarchy of ASs and areas (domains). Within areas, routers run interior gateway protocols (IGPs) such as OSPF or intermediate system-to-intermediate system (IS-IS) to maintain link state databases (LSDBs) [1]. The inter-AS level uses exterior gateway protocols (EGPs), most notably distance vector border gateway protocol (BGP), for reachability exchange. However, BGP represents a very high level of state aggregation and is generally insufficient for TE circuit routing, which requires link/path state [4]. Here, instead, OSPF-TE can provide an added inter-domain routing capability. However, with growing quality of service (QoS) requirements, OSPF has defined TE extensions (OSPF-TE, RFC 2676) for new opaque link state attributes (LSAs). Namely, these entities can disseminate QoS-related state to support advanced constraint-based routing (CBR). Note that QoS destination extensions also were proposed in BGP.

In recent years, the IETF also extended control provisioning to optical domains by augmenting its MPLS suite, namely, generalized MPLS (GMPLS) [1]. For routing, this includes new OSPF-TE opaque LSA definitions for DWDM and SONET/SDH links, enabling TE databases (TEDBs) to store wavelengths/usages, timeslots/usages, shared risk link groups (SRLGs), and so on. For signaling, extended Resource Reservation Protocol with Traffic Engineering (RSVP-TE) now supports hard state circuit set up/takedown, recovery, and so on. A new link management protocol (LMP) also is defined for resource discovery and fault localization. Now it is important to consider the applicability of GMPLS in multidomain settings. From a routing angle, OSPF-TE can suffice as a unified inter-domain link-state solution because it supports multiple link granularities (and is also being adopted in the OIF). RSVP-TE offers many saliences for multidomain circuit signaling via its loose route (LR) feature. Namely, partial skeleton routes can be specified, and subsequent explicit route (ER) expansion can be used to resolve full label switched paths (LSPs) [1]. RSVP-TE also defines mechanisms for LSP set up across domain boundaries — contiguous, stitched, and nested [3]. Carefully note that proxy devices also can be used to incorporate proprietary domains under the GMPLS framework, for example, see the network-aware resource broker (NARB) in a DRAGON network.

Another recent IETF multidomain standard is the path computation element (PCE) framework [4], which decouples TE path computation from signaling. In this setup, a domain can have one or more logical (standalone or co-located) PCE entities that communicate with path computation clients (PCC) to resolve connection paths. All PCC-PCE communication is performed via a new PCE protocol (PCEP) [4]. Although a PCE has access to local domain resource/policy databases, its inter-domain visibility may vary [3]. In the simplest case, a PCE may have knowledge only of its domain egress, namely, local visibility (low-trust, intra-carrier). Alternatively, a PCE may have knowledge of physical inter-domain links and even resources in external domains, namely, partial visibility (high-trust, intra-carrier). Accordingly, two distributed path computation schemes
are envisioned, per-domain and PCE-based [3]. The former computes paths in a domain-domain manner and is suitable for limited visibility. Namely, PCE entities (or border nodes) iteratively compute TE paths across their domains to ingress nodes in the next domain. The latter operates with increased inter-domain visibility, and two strategies have been presented: multi-PCE path computation with/without inter-PCE signaling [4], which will be discussed later. Note that the PCE framework also allows policy control at domain boundaries — a crucial requirement in multicarrier settings, on par with TE objectives. Specifically, an ingress PCE can enforce policies to determine the requests that it will support along with applicable TE constraints/algorithms.

**Research Survey**

Despite standards progress, the overall area of multidomain (multilayer) optical networking has not seen significant research focus. Although some results can be reported, most wireline multidomain efforts have focused largely on homogeneous packet networks. To get a better sense of the key challenges herein, it is important to survey the related areas, a generic taxonomy of which is shown in Fig. 2.

**Multidomain Packet-Switching Networks**

Many multidomain studies have looked at topology abstraction schemes for multidomain routing in IP and ATM networks, for example, [5, 6]. These schemes use graph transformations to condense resource state via virtual graphs with fewer abstract vertices and edges. Typically, this is performed by a designated domain-level entity; for example, an automatically switched optical network (ASON) routing controller (RC) [1], which propagates the abstract link state to other domains to build a global aggregated graph. Earlier studies in ATM networks using peer group summarization showed very good state reduction, order magnitude in nature. Subsequent efforts extended these concepts to IP networks as well. For example, [5] applies topology abstraction in multidomain IP QoS networks using star, mesh, tree, and spanner graphs. These schemes are tested with various path computation strategies (widest-shortest, shortest-distance) and show strong improvements in routing scalability and route fluctuation reduction. Further work in [6] studied aggregation in directed graphs with delay and bandwidth metrics using information-theoretic and line segmentation techniques. Overall results here showed good gains with aggregation, that is, higher success, lower crankback, and so on.

Apart from routing, other efforts also studied signaling crankback in multidomain IP/MPLS networks. For example, recently [3] detailed a compute while switching (CWS) scheme in which per-domain computation is used to set up an initial feasible route. Although data transmission is started on this initial route, crankback is used to search for more optimal routes (requiring new RSVP-TE attributes). Results show very high setup success, on a par with global state. Others also have looked at multidomain survivability. For example, most early schemes used per-domain primary/back-up LSP routing using BGP routing tables. However, this approach is very suboptimal and can easily overload heavily-traversed interdomain links. To address these limitations, [7] develops a more advanced scheme to capture domain-level diversity. Commensurate dedicated protection schemes also are developed using Surballe’s algorithm for trap topologies. However, generated state is quadratic in nature, that is, \(O(N^2)\) for \(N\) border nodes, and must be flooded at the inter-domain level. Moreover, detailed performance results are not presented. In [8] multidomain shared-path protection is studied, namely, an aggregated graph is defined with (full-mesh) virtu-
Multi-domain grooming pertains to the interconnection of networks with heterogeneous link granularities and can be considered as a subset of a multi-domain optical networking problem. This is a rather challenging topic with only a handful of studies conducted to date.

Multi-domain DWDM Networks

Multi-domain DWDM lightpath provisioning is an important focus area today. However, most related efforts have considered only next-hop domain routing approaches. For example, [9] details a domain-by-domain RWA scheme where gateways maintain complete alternate routes across all-optical and opto-electronic networks. Simulations show the overall effectiveness of this approach, although path dissemination is not studied. In [10], the authors study RWA for multisegmentation DWDM networks and develop three schemes, namely, end-to-end (E2E), concatenated shortest path (CSP), and hierarchical routing (HIR) are studied. Namely, the E2E scheme assumes a flat globalized graph; the HIR scheme assumes a hierarchical graph with segments summarized at nodes; and the CSP scheme simply uses local information for segment-by-segment routing. Results for a specialized mesh-torus topology show significant blocking reduction with the E2E and CSP schemes. However, no intra or inter-domain routing is performed here.

Recent proposals also extended topology abstraction to multi-domain DWDM settings, where optical links pose very different constraints from basic delay and bandwidth metrics, for example, wavelengths, timeslots, converters, SRLGs, and so on. In [11], a theoretical study of partial information models for domains with border node conversion is presented. Lightpath selection is treated as a Bayesian decision, and findings show that scalable information models (i.e., logarithmic growth per wavelengths) achieve a good trade-off with loss (Bayes error rate). However, this treatment is theoretical and focuses on bus topologies. Also, inter-domain routing and RWA are not studied. In [12], a simple-node abstraction scheme for ASON is tabled. However, signaling and wavelength conversion are not considered — the latter being a key necessity at domain boundaries performing regeneration and bit-level service level agreement (SLA) monitoring. Finally, [14] develops full mesh and star abstractions for all-optical and opto-electronic networks using modified RWA schemes. Inter-domain update strategies and distributed RWA also are designed. Overall findings show much-improved blocking performance (and lowest signaling load) with full mesh abstraction; albeit routing overheads increase with the square of border nodes, namely, $O(N^2)$ mesh.

Note that DWDM sub-path protection is also a well-studied topic and bears some relevance to multi-domain RWA. Namely, an inter-domain lightpath can be segmented ideally into protection domains and various sub-path/segment protection schemes applied therein, for example, dedicated, shared, and so on. However, most DWDM sub-path schemes are premised upon the availability of global state information and will require significant adaptations for distributed multi-domain operation.

Multi-domain Grooming

Multi-domain grooming pertains to the interconnection of networks with heterogeneous link granularities and can be considered as a subset of a multi-domain optical networking problem. This is a rather challenging topic with only a handful of studies conducted to date. For example, [13] considers multi-domain provisioning in SONET-DWDM networks and defines a segment graph model to evaluate three path selection schemes — centralized (full-knowledge), domain-by-domain (local knowledge), and hierarchical source routing (partial inter-domain knowledge). The latter propagates domain state only for specified granularities, but state aggregation is not detailed. The results show much lower blocking with increased levels of inter-domain state propagation. More recently, the authors in [15] developed a novel framework for direct Ethernet-DWDM integration using integrated multi-domain signaling to simultaneously provision lightpaths and groom sub-rate EPL circuits. However, underlying DWDM networks are assumed to static and routers perform only at the Ethernet layer. Detailed performance results with varying routing update strategies and crankback show minimal increases in set-up latencies and signaling loads.

Open Research Challenges

By and large, the study of multi-domain/multi-granularity optical networks is in its infancy. A simple solution to this problem can be to use static partitioning. For example, SONET/SDH or NGS/MSP domains can be interconnected via pre-configured multi-domain lightpaths, based upon integer linear programming (ILP) optimizations. Here a full range of existing (single-domain) algorithms can be reapplied, for example, RWA, grooming, survivability, and so on. Because these partitions are dedicated, there also are no interactions between the respective control planes, for example, separate routing, signaling instances. Although feasible, this approach is very inefficient for dynamic demands and requires constant reoptimization. Hence, it is not considered further herein and instead, more viable integrated schemes are discussed. The major functional areas here include multilayer state dissemination, TE path computation/signaling, and survivability (Table 2).

State Dissemination

Various inter-domain dissemination models are possible with the simplest performing no state exchange across boundaries, that is, local visibility [3]. In this case, a domain is simply abstracted as a simple node; see Fig. 3. However, because inter-domain TE link/path state is very beneficial for multi-domain networks, ideally partial visibility is desirable. For example, this can entail distribution of physical interdomain link state (SONET/SDH, DWDM) or aggregated domain state. The net result is a global multi-granularity network view comprising physical and abstract links, namely, aggregated or abstract
graphs [14]; see Fig. 3. Note that here inter-domain TEDBs must maintain multiple link types, for example, via different OSPF-TE opaque LSA types. Nevertheless, maintaining partial visibility presents several challenges. Foremost, it mandates the design of effective resource aggregation schemes to support TE objectives (as these directly impact the trade-off between routing scalability and blocking). A promising approach is to leverage the topology abstractions developed for IP, and more recently DWDM, networks, as shown in the aggregated graph view in Fig. 3. However, specific considerations are required for heterogeneous optical domains, particularly for multiple switching granularities. Additional efforts also must look at summarizing survivability state, such as SRLG, fiber/span protection, and so on. Finally, the aggregation of inter-domain lightpath connections also can be performed as these entities can be treated as virtual connection grooming links (i.e., hierarchical LSP) between SONET/SDH domains; see Fig. 3. Although this step can boost grooming efficiencies, very scalable algorithms are required because connection counts can grow with the square of the nodes. Note that policy constraints play a key role in determining what TE constraints and aggregation schemes are allowed across domain boundaries. Also, re-optimization (traffic re-routing) may be required to further maintain efficiency on such virtual connection links.

Finally, resource dissemination requires effective inter-domain update triggering strategies [17]. Various routing policies are available in the aggregated graph, such as timer-based, absolute change, relative change, and hysteresis-based. All of these schemes — except for the former — are threshold-based, and extensive findings show that such strategies yield the lowest routing overheads and good sensitivity [5]. Hence, physical inter-domain link state can readily be propagated accordingly. Conversely, aggregated link state propagation is more complicated as it pertains to “non-existing” computed entities. Here, using fixed intervals to compute abstractions and propagate updates can yield high messaging/compute loads (short intervals) or high inaccuracy (long intervals). Ideally, hybrid periodic/threshold-based strategies must be developed [14, 17].

### TE Path Computation and Signaling

The main objective of path computation is to set up feasible end-to-end circuit routes that meet required TE and policy objectives, preferably with some degree of global optimality. Given the distributed nature of the problem, requisite signaling — RSVP-TE and PCEP — will play a crucial role. Herein, two broad strategies are possible for optical domains, per-domain and PCE-based computation [3].

**Per-Domain Strategies** — Per-domain computation is suited for low inter-domain visibility scenarios. However, it is very difficult to select a new top domain (in pursuit of TE objectives) in a local manner. As such, the only viable option may be to choose predetermined or random egress points, yielding low resource efficiency/high blocking, particularly for multigranularity grooming. To address these limitations, crankback signaling can be used to reattempt setups at intermediate domain boundaries. The trade-off here is in terms of set-up delay and messaging. Recent findings in multidomain MPLS networks show that crankback gives significantly lower blocking when using active (failed) path state [3]. Hence, further studies can develop multilayer optical crankback schemes to incorporate wavelength selection and even virtual connection (grooming) state. Periodic/event-driven reoptimization of connections (or connections groups) also was suggested in [4]; albeit in-depth analyses are required to devise effective algorithms such as graph, theoretic, or ILP schemes. Nevertheless, reoptimization can be very costly to implement across domains, particularly if clients demand hitless switchovers.

**PCE-Based Strategies** — As mentioned earlier, two PCE strategies are possible: multi-PCE computation with or without inter-PCE signaling. The latter assumes partial visibility and requires a source PCE to compute a skeleton LR to the destination domain. In practice, this is best performed via modified graph-theoretic searches running over the aggregated network graphs. However, detailed multidomain grooming schemes are lacking. To address this gap, two-step schemes can be considered, akin to single-domain grooming. Namely, an initial LR can be searched on the aggregated graph using only links with the same granularity as the source node. Various TE objectives can be studied here, including minimum hops/cost, load balancing, and so on [5]. Pending failure at this “higher level, subsequent virtual connection lightpaths can be set up between border nodes. However, it is not clear how to identify the best border node end-points; for example, [14] uses exhaustive search. Overall, PCE-based LR schemes will be more effective than basic per-domain computation, but more studies are required to assess the impact of aggregation on global optimality.
Survivability schemes can generally be classified as either pre-configured protection or post-fault restoration. The former are best suited for stringent demands, whereas the latter are more germane for delay-tolerant traffic.

Path computation with inter-PCE signaling requires a source PCE to request other PCE entities to resolve explicit or loose source-destination routes. This approach is amenable to low trust inter-carrier settings and mimics per-domain computation (except that signaling occurs on the PCE plane). Along these lines, a specialized backward recursive PCE-based computation (BRPC) was presented for IP/MPLS PCE-networks [16], expanding strict hops backwards from the destination domain using a pre-defined virtual shortest path tree (VSPT). This set up assumes a-priori specification of end-to-end domain sequences and extensions to PCEP messaging. However, detailed analyses and/or DWDM network adaptations are lacking as of now.

Note that PCE-based computation requires follow-up RSVP-TE signaling, which in turn requires considerations for differing link types. For example, ingress border nodes must resolve LR hops (abstract links) with traversing intra-domain sequences. Metric translation at domain boundaries may be required here because domain TE objectives will vary versus global inter-domain TE objectives. Furthermore, if an ingress border node runs at a lower granularity than the incoming request, for example, SONET over DWDM, grooming will be required over existing and/or new intra-domain lightpaths, namely, hierarchical/nested LSP mapping. In this case, a wide range of existing (intra-domain) grooming schemes can be reapplied.

**Survivability**

Survivability schemes can generally be classified as either pre-configured protection or post-fault restoration. The former are best suited for stringent demands, whereas the latter are more germane for delay-tolerant traffic. In extending these paradigms to heterogeneous optical settings, a simplifying assumption can be to run only inter-domain restoration routes to the same domain sequences. Hence, intra-domain failures can be resolved using robust, these schemes are costly and resource inefficient as they relegate primary and backup routes to the same domain sequences. New protection strategies with domain diversity can be developed to improve resiliency, using either per-domain or PCE-based computation. These schemes can be further coupled with sequential or simultaneous (parallel) set-up strategies [16]. For example, sequential per-domain protection can first resolve primary paths and use the returned explicit routes to avoid overlaps on the back-up paths, for example, via the RSVP-TE explicit, record, and exclude route objects (ERO, RRO, and XRO). However, this likely will yield sub-optimal routes and higher blocking (also susceptible to “trap” topologies). To address these concerns, again, crankback and reoptimization can be performed [4]. Note that inter-carrier confidentiality can restrict the sharing of explicit routes between domains. Workarounds can include the use of path keys indexing or pre-established traversing paths [16].
In PCE-based protection computation (without inter-PCE signaling), source PCE entities must compute disjoint path pairs as per TE objectives, for example, minimizing total cost. One approach can be to compute domain-disjoint primary/back-up LR sequences over the aggregated graph. Nevertheless, this solution is quite restrictive and will yield high blocking, especially with low inter-domain connectivity. Although more advanced schemes can relax this requirement and allow domain overlaps between primary/back-up routes, ensuring intra-domain diversity becomes a concern. For the case of sequential set ups, the implications will be similar to those discussed for per-domain protection. Conversely, for simultaneous set ups, specialized aggregated graph state will be required to ensure intra-domain diversity; see [7, 8]. Note that there is also a grooming aspect to multidomain optical protection. Herein, recent studies on survivable SONET-DWDM grooming looked at protection-at-lightpath (PAL) and protection-at-connection (PAC) concepts. The overall findings here show much-improved efficiency and scalability with lower-layer PAL. However, more work is required to extend these concepts to multiple domains with abstracted diversity state.

**Restoration Schemes** — Restoration uses post-fault signaling to re-route failed circuits across multiple domains. In essence, this approach has much commonality with per-domain computation with crankback. Hence, similar strategies can be developed to couple failure/fault state with RSVP-TE XRO objects to drive circuit re-routing. Multilevel SLA support can also be considered, with higher priority connections preempting lower priority resources during restoration crankback. Nevertheless, much work remains to analyze these schemes and their trade-offs in optical settings.

**Sample Study**

Recent work on multidomain DWDM provisioning is presented in [14] and incorporates many of the key components of a multilayer solution. Foremost, hierarchical routing is performed using two-level OSPF-TE to disseminate physical inter-domain link state (simple node abstraction). Optionally, aggregated domain state also is shared between border OXC nodes via full-mesh abstraction. Specifically, novel k-shortest path RWA schemes are developed to condense wavelength/converter resources along routes between border nodes and generate abstract links. The propagated link state then is used by source PCE entities to build aggregated DWDM graphs for LR sequence computation (minimum hop or load-balancing). Then, these lightpaths are signaled and expanded using RSVP-TE with wavelength trace vectors. Because routing scalability is a major concern for full-mesh abstraction, two improved update strategies also are defined herein: maximum relative change (MRC) and most-used maximum relative change (MU-MRC) [17]. The former limits updates to abstract links experiencing maximum wavelength changes, up to $h$ updates per domain; whereas, the latter further restricts updates contingent to the actual wavelengths that change status (and not just the number, as in MRC). In other words, only those abstract links experiencing the most number of changes in the MU wavelengths are updated, that is, the top $m$ out of top $h$; see [17] for details.

The previous schemes are tested using OPNET Modeler™ for a sample nine-domain network with 19 bidirectional inter-domain links; see [14].
All connections are generated between random domains using a 70-30 percent intra/inter-domain ratio, chosen to reflect practical settings. Furthermore, OSPF-TE hold-down timers are set to 100 s and relative update thresholds to 10 percent. Each run is averaged over 200,000 connections, and mean holding times are set to 600 s (exponential). First, inter-domain RWA blocking is plotted in Fig. 4a for no conversion (all-optical), sparse conversion (border nodes only), and full conversion (all nodes) with 16 wavelengths/link. In the latter cases, opto-electronic nodes are assigned 10 converters each in a share-per-node architecture. These results clearly show that full-mesh abstraction gives excellent blocking reduction versus no abstraction (simple node); for example, full-mesh blocking is over 40 percent lower for full conversion networks at 90 Erlang. Full mesh abstraction with sparse conversion also outperforms no abstraction with full conversion. This indicates that inter-domain routing can effectively lower hardware converter counts, enabling carriers to concentrate these costly devices at domain boundaries where they are most needed for regeneration/monitoring. The fraction of signaling failures (out of total failed set ups) also are plotted in Fig. 4b and Fig. 4c. Again, full-mesh abstraction yields the lowest false-start signaling rates, reducing control plane processing and signaling overheads.

Finally, the two improved update triggering policies are tested for full wavelength conversion and full-mesh abstraction. First, inter-domain blocking is plotted in Fig. 5a and indicates nearly identical performances between the regular significance change factor (SCF)-based scheme and the newer MRC and MU-MRC strategies. Next, Fig. 5b shows the inter-domain OSPF-TE routing loads (LSA/s) and confirms sizeable reductions, for example, almost 50 percent lower than SCF-based updates at light-to-moderate loads. Overall, these results confirm that improved DWDM update policies can boost inter-domain routing scalability without compromising blocking. In particular, the MU-MRC scheme gives the best performance, further indicating the efficacy of inter-domain MU state.

CONCLUSIONS

This article surveys multidomain/multilayer optical network control. First, relevant standards are reviewed and their applicability to multidomain operation discussed. Then, the existing body of research is presented and challenges are identified in the areas of state dissemination, TE path computation, and survivability. Finally, a sample multidomain DWDM study is presented, with overall findings confirming the effectiveness of scalable state aggregation and update strategies in lowering blocking.

ACKNOWLEDGEMENTS

This research was supported in part by the U.S. Department of Energy (DOE) Office of Science under award number ER25828 and the U.S. National Science Foundation (NSF) Computer and Information Sciences Division (CISE) under award number CNS0806637. The authors are very grateful to these agencies for their generous support. In addition, the authors are also very thankful to Dr. Thomas Ndousse for his insightful discussions and feedback.

REFERENCES

and IEEE GLOBECOM 2006. He is also a program committee, industry magazine articles, and standardization applications (including journal and conference papers, book Networks, and Tennessee Tech University. He is a recipient has held senior positions at Nokia, IBM, Motorola, Sorrento funded research projects. He gained a wide range of indus- provided Ph.D. degree in the Department of Electrical and Computer Engineering at the University of New Mexico, Albuquerque. In the past he has worked as a summer intern at Motorola China and Acadia Optronics LLC, USA, and as a software engineer at Bell Labs/Lucent Technologies China. He has authored or co-authored over 15 publications, and his research interests include optical networks, multidomain networks, traffic engineering, multihop communications, and wireless networks. ASHWIN GUMASTE is currently a faculty member in the Department of Computer Science and Engineering at the Institute of Technology, Bombay, and a visiting scholar at the Massachusetts Institute of Technology. He was previously with Fujitsu Laboratories Inc. as a member of the research staff in the Photonics Networking Research Lab (2001–2005). Prior to this he worked in Fujitsu Network Communications R&D and prior to that, with Cisco Systems in the Optical Networking Group (ONG). He has over 40 pending U.S. and EU patents, and has published close to 60 articles in referred conferences and journals. He has authored three books on broadband networks: DWDM Network Designs and Engineering Solutions (a networking bestseller). First-Mile Access Networks and Enabling Technologies (Pearson Education/Cisco Press), and Broadband Services: User Needs, Business Models and Technologies (Wiley). He is also an active consultant to industry, and has worked with both service providers and vendors. In addition, he has served as program chair, co-chair, publicity chair, and workshop chair for various IEEE conferences, and has been a technical program committee member for ITCC, ICC, GLOBECOM, IEEE Broadnets, IEEE ICCCN, Gridnets, and so on. He is also a guest editor for IEEE Communications Magazine and was General Chair of the 1st International Symposium on Optical Networks and Telecommunication Systems, held in Bombay, India. DISS BENHADDOU received a Bachelor of Science in Physics (electronics) from the Abdelmalek Essaadi, Morocco. He received a Master of Science in optoelectronics from the University of Montpellier II, France, in 1991. He holds two doctoral degrees in optoelectronics and telecommunications engineering from the University of Montpellier II in 1995, and the University of Missouri in 2002. He is an assistant professor at the University of Houston (UH), Texas, where he is actively involved in optical networking and sensor network research activities. Prior to joining UH, he was a senior technical staff member at Lambda Optical Systems Inc., where he played a key role in protocol development and systems integration activities. In particular, he led systems test/integration activities for the Advanced Technology Demonstration Network (ATDNet) testbed project and worked closely with the Naval Research Laboratory (NRL) and the Laboratory for Telecommunication Sciences (LTS). During his earlier tenure at Sprint, he also implemented an extensive broadband testbed for vendor equipment certification and research/development activities. NAGESWARA S. V. RAO [F] received a B.Tech from the National Institute of Technology, Warangal, India in 1982, an M.E. from the School of Automation, Indian Institute of Science, Bangalore in 1984, and a Ph.D. in computer science from Louisiana State University in 1988. He is currently UT-Battelle Corporate Fellow at Oak Ridge National Laboratory, which he joined in 1993. He was also an assistant professor of computer science at Old Dominion University from 1988 to 1993. His research interests include high-performance networking, network transport dynamics, and information and sensor fusion. He has published more than 250 technical journal and conference papers. He is an Associate Editor for Pattern Recognition Letters, International Journal of Distributed Sensor Networks, and Information Fusion. He received the 2005 IEEE Computer Society Technical Achievement Award for his contributions in the information fusion area.

BIographies

NASIR GHANI [S’95, M’97, SM’01] (nghani@ece.unr.edu) received a Bachelor’s degree in computer engineering from the University of British Columbia, Canada, in 1992, and a Master’s degree in electrical engineering from McMaster University, Canada, in 1992, and a Ph.D. degree in electrical and computer engineering (ECE) from the University of Waterloo in 1997. Currently, he is an associate professor in ECE at the University of New Mexico, where he is involved in various funded research projects. He gained a wide range of industrial and academic experience in the networking area, and has held senior positions at Nokia, IBM, Motorola, Sorrento Networks, and Tennessee Tech University. He is a recipient of the NSF Career Award and has authored over 100 publications (including journal and conference papers, book chapters, industry magazine articles, and standardization drafts), as well as two patents. He is a co-chair of the IEEE INFOCOM High-Speed Network (HSN) series and has served as a co-chair for the optical symposia for IEEE ICC 2006 and IEEE GLOBECOM 2006. He is also a program committee member for IEEE/OSA OFC and numerous other conferences, and has served regularly on NSF, DOE, and international panels. He is an associate editor of IEEE Communications Letters and has guest edited IEEE Network, IEEE Communications Magazine, and Cluster Computing.

QING LIU received a B.E. degree in telecommunication engineering and an M.E. degree in computer science from Nanjing University of Post and Telecommunications, China, in 2001 and 2004, respectively. He is currently pursuing a Ph.D. degree in the Department of Electrical and Computer Engineering at the University of New Mexico, Albuquerque. In the past he has worked as a summer intern at Motorola China and IEEE Communications Magazine • June 2008