Available online at www.sciencedirect.com
ScienceDirect
Computer Communications 30 (2007) 3662–3675
computer communications
www.elsevier.com/locate/comcom

Distributed inter-domain lightpath provisioning in the presence of wavelength conversion

Q. Liu [a], N. Ghani [a,*], N.S.V. Rao [b], A. Gumaste [c], M.L. Garcia [d]

[a] *Department of Electrical & Computer Engineering, University of New Mexico, Albuquerque, NM 87131-0001, USA*
[b] *Computer Science and Mathematics Division, Oak Ridge National Laboratory*
[c] *School of Information Technology, IIT Bombay*
[d] *ECE Department, Tennessee Tech University*

Received 10 February 2007; received in revised form 24 June 2007; accepted 28 July 2007
Available online 10 August 2007

Abstract

Dense wavelength division multiplexing (DWDM) has become the dominant transport layer technology for next-generation backbone networks due to its unprecedented capacity scalability. As a result, there is a pressing need to investigate lightpath provisioning in multi-domain DWDM networks. Although inter-domain provisioning has been well-studied for packet/cell-switching networks, the wavelength dimension (along with wavelength conversion) presents many added challenges. To address these concerns, a detailed GMPLS-based hierarchical routing framework for provisioning transparent/translucent/opaque multi-domain DWDM networks is presented. The scheme adapts topology abstraction to hide internal domain state so as to resolve routing scalability and security issues. Specifically a novel full-mesh topology abstraction scheme is developed for full wavelength conversion domains, i.e., to disseminate additional wavelength converter state. Related inter-domain lightpath RWA and signaling schemes are also tabled. Performance analysis results are then presented to demonstrate the effectiveness of the proposed mechanisms along with directions for future research work.

Keywords: Multi-domain networks; Optical networks; GMPLS; Inter-domain DWDM routing; Topology abstraction; Wavelength conversion

1. Introduction

To date, DWDM has emerged as the premiere transport technology and has gained much traction in long-haul and metro/regional domains [1]. DWDM exploits the huge unused spectrum in *single mode fiber* (SMF) to transmit multiple channels at unprecedented terabits/s speeds. As this technology has matured, a wide range of circuit-switching capabilities have evolved, e.g., *optical cross-connect* (OXC) and *optical add-drop multiplexer* (OADM) devices. Additionally, there has been much progress in architectures and frameworks for optical networks. Namely, the IETF generalized *multi-protocol label switching* (GMPLS) framework [2] has adapted packet-based *multi-protocol label switching* (MPLS) protocols for provisioning "non-packet" circuit-switched connections, i.e., via label abstractions for wavelengths, timeslots, etc. GMPLS includes key additions for routing, signaling, and link discovery [2]. Concurrently, the ITU-T has also specified a broad-based *automatic switched optical network* (ASON) framework as well.

With regards to provisioning algorithms, a multitude of DWDM *routing and wavelength assignment* (RWA) [3] and survivability schemes have been evolved. Multi-layer grooming schemes between DWDM and SONET or IP networks have also been addressed in detail, see survey in [1]. However most of these efforts have focused on *single* domain networks. As DWDM technology proliferates there is now a pressing need to develop more advanced lightpath provisioning algorithms for distributed *multi-domain* settings, particularly those with wavelength conversion. In particular emergent applications in the fields of

* Corresponding author. Tel.: +1 505 277 2436; fax: +1 505 277 1439.
E-mail address: nghani@ece.unm.edu (N. Ghani).

0140-3664/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2007.07.006

grid-computing and e-science are driving the need for distributed, dynamic circuit-switched interconnection at very large speeds.

As the number of DWDM domains increases it becomes difficult for a single entity to maintain detailed *global* state across all domains, e.g., physical links, available resources, diversity, etc. Indeed it is evident that some form of information *aggregation* and distribution is necessary. Furthermore there is also a commensurate need for inter-domain lightpath RWA algorithms that leverage this aggregated state. This can be achieved via hierarchical routing and topology abstraction techniques which have been well-studied for packet-switching IP and/or cell-switching *asynchronous transfer mode* (ATM) networks [5,6]. For example the ATM Forum's *private network to network interface* (PNNI) protocol [1] clusters nodes into peer groups and uses topology abstraction to hide internal state from outside users [5]. Although these principles have recently been extended for DWDM networks [7,17], these studies have not considered the impact of large scale wavelength conversion. Given the increased distance and coverage of such networks, it is indeed likely that wavelength conversion and regeneration will be a vital necessity.

This paper addresses the important area of distributed *inter-domain* provisioning in full wavelength conversion DWDM networks and builds upon the initial contribution in [7]. Namely, a two-level hierarchical routing model is developed using the GMPLS framework for full wavelength conversion multi-domain networks. This solution defines two topology abstraction algorithms and also tables associated inter-domain lightpath RWA and signaling schemes. The paper is organized as followings. Section 2 presents a background review. Subsequently, Section 3 details the proposed inter-domain routing framework whereas Section 4 focuses on commensurate inter-domain lightpath provisioning and signaling algorithms. Next, simulation results are presented in Section 5 along with conclusions and future work in Section 6.

2. Background

Multi-domain circuit-switched DWDM networks have received notable attention from a standards perspective. For example the ASON framework defines a comprehensive multi-level routing hierarchy consisting of different areas [2]. However detailed protocols/algorithms are not specified here. Meanwhile the OIF has also evolved a generic *network-to-network* (NNI) specification [2,5] to facilitate multi-domain provisioning, e.g., reachability/resource exchange, setup signaling, etc. Various IETF proposals have also tabled *traffic engineering* extensions to inter-area protocols [8] and inter-domain *border gateway protocol* (BGP) [9] to propagate resource state information. For example, St. Arnaud [9] proposes route advertise/withdraw messaging between domains using modified BGP along with proxy *lightpath route arbiters* (LRA) to compute routes between border OXC nodes. Nevertheless the detailed *algorithmic* study of inter-domain DWDM routing/provisioning algorithms has not been considered and this is only now gaining attention.

Various other inter-domain DWDM studies have been presented in [10–12,18], with a primary focus on lightpath computation and signaling. For example, Yang and Ramamurthy [10] detail a hop-by-hop routing and signaling scheme in which domain gateways maintain complete (alternate) route state. However related resource propagation issues (i.e., path dissemination) are not considered and hence this setup is more favorable to BGP-type implementations. Meanwhile [11,16] table a hierarchical inter-domain solution for ASON based upon a simple node (coarse) abstraction policy. However no signaling provisions are considered here and hence the scheme is most amenable to centralized implementation. Additionally, opto-electronic conversion is not considered – a vital necessity in larger inter-domain settings. Additionally, Liu et al. [18] present a theoretical treatment of state compression in multi-domain DWDM networks with border node conversion. Here, various information models are studied and lightpath selection is modeled as a Bayesian decision problem. However, this treatment is largely theoretical and the results only consider bus topologies. Furthermore, no routing protocols or distributed lightpath setup are detailed. Note that others have also considered sub-path self-healing loops (domain) protection strategies for optical lightpaths, e.g., [12]. Although these schemes present many saliencies, they premised upon the availability of global state, i.e., "flat/single-domain" network. Hence the extension of these schemes into distributed multi-domain DWDM networks with no/limited global state is not straightforward. A distributed framework is now presented to address some of these challenges.

3. Inter-domain routing

Distributed inter-domain DWDM provisioning is a challenging problem as the wavelength/converter dimension poses many restrictions. Herein a comprehensive framework is developed for *distributed* multi-domain lightpath RWA in both all-optical *and* hybrid opto-electronic networks. In the latter it is assumed that only the border nodes are conversion-capable and the interior nodes remain all-optical. This is a realistic modeling of emergent optical networks, e.g., all-optical "islands" delineated with optoelectronic border nodes [4]. The proposed scheme addresses several key steps in inter-domain provisioning. Foremost, a multi-domain *topology abstraction* model is defined to condense domain-level DWDM state. Subsequently, appropriate *inter-domain routing and triggering policies* are developed to disseminate both physical/abstracted inter-domain state. Note that this approach favors link-state routing implementations which are most suitable for the increased dimensionalities of DWDM networks.

Topology abstraction is used to summarize domain-level state. In particular the proposed hierarchical scheme designates a specific *border* OXC node in each domain as a *routing area leader* (RAL) [2]. This entity computes a DWDM topology abstraction for its domain by transforming the physical topology into a virtual topology. Specifically, two abstraction schemes are presented, i.e., simple node and full-mesh (Fig. 1), as evolved from earlier proposals for data/cell-switching networks [5]. Note that more advanced star abstractions are left for future study, see [17] as well. The *virtual* state information (as generated via the abstractions) is then flooded to border OXC nodes across all domains using inter-domain routing protocols. These helps maintain a synchronized global virtual view of the whole network and allows abstracted information to be used to compute end-to-end inter-domain lightpaths.

To detail the topology abstraction schemes it is first necessary to develop the required notation. Here all set and vector entities are denoted in bold and it is assumed (without loss of generality) that fiber connectivity is bi-directional, i.e., there are two opposite-direction fiber links between a pair of connected OXC nodes. Consider a DWDM network comprising D domains, with the i-th domain having n^i nodes and b^i border/gateway OXC nodes, $1 \leqslant i \leqslant D$. This network is modeled as a collection of domain sub-graphs, $G^i(\boldsymbol{V}^i, L^i)$, $1 \leqslant i \leqslant D$, where $\mathbf{V}^i = \{v_1^i, \ldots, v_{n^i}^i\}$ is the set of physical domain OXC nodes and $\boldsymbol{L}^i = \{l_{jk}^i\}$ is the set of physical *intra-domain* links in domain i ($1 \leqslant i \leqslant D$, $1 \leqslant j'$, $k \leqslant n^i$), i.e., l_{jk}^i is the link between OXC nodes v_j^i and v_k^i. All links have W wavelengths. Furthermore, $\boldsymbol{B}^i$ represents the set of border OXC nodes within domain i and without loss of generality, it is assumed that these nodes are numbered as the *first* group of nodes in the domain, i.e., $\boldsymbol{B}^i = \{v_1^i, \ldots, v_{b^i}^i\}$. Meanwhile for multi-domain routing, a *higher-level* topology is defined containing the border OXC nodes and inter-domain links. This is given by a graph $H(\boldsymbol{U}, \boldsymbol{E})$, where $U = \sum_i \{\boldsymbol{B}^i\}$ is the set of border OXC nodes across all domains ($1 \leqslant i \leqslant D$) and $\mathbf{E} = \{e_{km}^{ij}\}$ is the set of physical *inter-domain* links, i.e., e_{km}^{ij} inter-connects v_k^i in domain i with v_m^j in domain j, $1 \leqslant i, j \leqslant D$, $1 \leqslant k \leqslant b^i$, $1 \leqslant m \leqslant b^i$. This graph contains all physical border nodes and inter-domain links but does not necessarily have full connectivity – which is achieved via appropriate topology abstractions (detailed next). Note that since DWDM networks are being treated, all links (physical, virtual) have associated binary *wavelength availability vectors*, i.e., λ_{km}^{ij} is the wavelength availability vector for link e_{km}^{ij}, where $\lambda_{km}^{ij}(n) = 1$ if the n-th wavelength is available and 0 otherwise, $1 \leqslant n \leqslant W$. Two different abstraction schemes are now presented.

3.1. Simple node (SN) abstraction

This is the simplest of all the abstraction schemes and condenses a domain into a single *virtual* node emanating all physical inter-domain links, as proposed in [7]. For

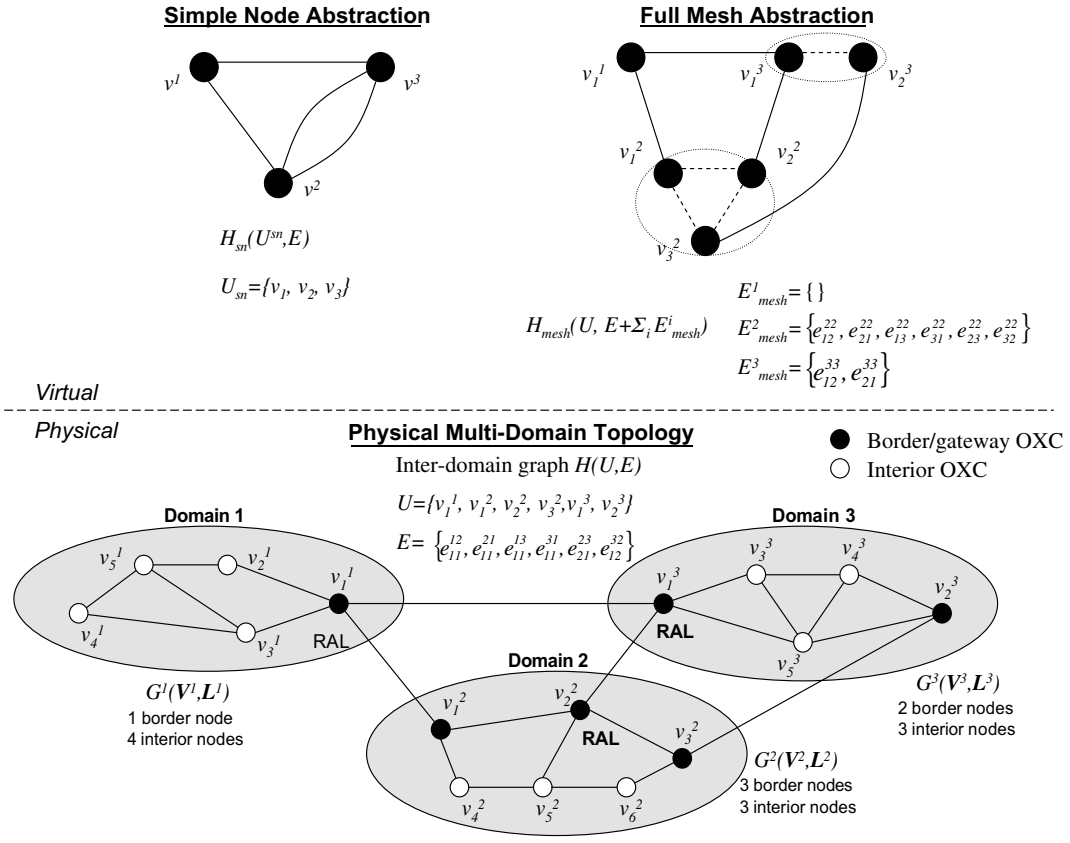

Fig. 1. Simple node and full-mesh topology abstraction.

example, the three border OXC nodes in domain 2 in Fig. 1 are simply collapsed into a single virtual node with 3 inter-domain links. This scheme provides no visibility into domain-internal state (i.e., wavelengths, converters) and has low inter-domain routing overheads. Mathematically, the above transformation can be represented as $H(\boldsymbol{U},\boldsymbol{E}) \rightarrow H_{sn}(\boldsymbol{U}_{sn},E)$ where $\boldsymbol{U}_{sn} = \{v^i\}$ is the condensed set of virtual nodes representing each domain i and $\boldsymbol{E}$ is the set of physical inter-domain links. Namely, border node set $\boldsymbol{B}^i$ is mapped to a single virtual node vertice, v^i, e.g., see Fig. 1. Overall the simple node scheme requires state for all physical inter-domain links, i.e., order O($|\boldsymbol{E}|$) storage/update complexity.

3.2. Full-mesh abstraction

The full-mesh abstraction scheme is designed to perform intra-domain state summarization involving per-link wavelength state and possibly per-node converter state. Herein several wavelength conversion scenarios are treated, namely *no conversion* (i.e., transparent or "all-optical"), *sparse wavelength conversion*, and *full wavelength conversion*, respectively. The former scheme assumes all-optical networks and is the same as the similarly titled scheme in [7]. The latter two schemes assume some level of wavelength conversion (i.e., signal regeneration) within the network. In particular, sparse wavelength conversion assumes that only border nodes have (a limited number of) wavelength converters and these are strictly reserved for inter-domain lightpath provisioning only. As such this is the same as the optoelectronic scheme presented in [7]. However, the latter scheme (i.e., full wavelength conversion) is a new extension proposed herein to address the more realistic case of wavelength conversion at interior nodes as well. Namely, regular intra-domain lightpaths can also undergo wavelength conversion. Detailed full-mesh abstraction algorithms are now presented for each of these cases.

3.2.1. No conversion and sparse conversion

In both of these cases, the interior OXC nodes do not support any wavelength conversion capability. As such only wavelength state has to be summarized and divulged to other domains, i.e., no interior wavelength converter state to be handled. Therefore both scenarios can be treated in an identical manner. Namely, the i-th domain $G^i(\boldsymbol{V}^i,\boldsymbol{L}^i)$ is transformed to a sub-graph containing the border nodes interconnected via a fully meshed set of *virtual* links, i.e., $\boldsymbol{E}^i_{\text{mesh}} = \{e^{ii}_{jk}\}$. Specifically, *wavelength availability vectors* are computed for all of these virtual links, i.e., λ^{ii}_{jk}, to summarize the wavelength state needed to traverse through the domain between border OXC nodes. For example, sub-graph $G^2(\boldsymbol{V}^2,\boldsymbol{L}^2)$ in Fig. 1 is transformed by connecting the border nodes with virtual links e^{22}_{12}, e^{22}_{21}, e^{22}_{13}, e^{22}_{31}, e^{22}_{23}, e^{22}_{32}. This abstraction can be represented by the following transformation $H(\boldsymbol{U},\boldsymbol{E}) \rightarrow H_{\text{mesh}}(\boldsymbol{U},\boldsymbol{E} + \sum i\{\boldsymbol{E}^i_{\text{mesh}}\})$, where $\boldsymbol{E}^i_{\text{mesh}}$ is the above-computed set of domain-level virtual links ($1 \leqslant i \leqslant D$). Overall the computation of virtual link wavelength availability vectors uses the K-shortest path and vector AND-ing techniques (full details of which can be found in [7]).

3.2.2. Full conversion

Full wavelength conversion networks pose much more complexity for full mesh abstraction since two resource dimensions must be summarized, i.e., wavelength and converter. Namely, in this case, both interior and border node can have a pool of converters. Now consider the fact that there are several wavelength converter node models, e.g., share-per-link and share-per-node model [15]. In order to economize converters usage and simplify node design, the share-per-node model is adopted in this work, although share-per-link can also be incorporated. Namely, converters are shared among all physical output links on an OXC node.

Now consider full-mesh abstraction with interior node conversion. Simply advertising the number of free converters along a virtual link between ingress/egress border node is not feasible. Namely, one must also convey wavelength converter locations and check the wavelength continuity constraint on each "sub-path", i.e., at least one common wavelength has to be available between two adjacent conversion-enabled OXC nodes. To address this challenge a novel full-mesh topology abstraction is developed, as shown in Fig. 2. The idea behind this scheme is to leverage path computation techniques to capture (i.e., transform) wavelength converter state into the actual wavelength availability vectors of the virtual links. This precludes having to decouple and separately advertise (domain-internal) wavelength and converter state. Fig. 3 shows an example of wavelength conversion inside domain i. Here the most congested *sub-path* will represent the traversing cost between border nodes v^i_1 and v^i_2, e.g., path v^i_1-v^i_3-v^i_4-v^i_2 becomes a feasible candidate if conversion is enabled at node v^i_3. Moreover, sub-path v^i_3-v^i_4-v^i_2 with wavelength availability vector [101000] will be the bottleneck segment and is therefore taken to represent the cost of the whole path, i.e., the virtual link between border nodes v^i_1 and v^i_2. In essence, this bottleneck *sub-path* state indirectly captures the wavelength conversion capabilities of the domain between the two respective border nodes.

Now consider the details of the full-mesh abstraction/graph transformation algorithm in Fig. 2. Here the scheme loops through each border node pair (indices j, k) and computes the associated wavelength availability vector for the corresponding virtual link. The scheme first runs the K-shortest path algorithm to generate a set of paths between each border node pair, denoted as $\{\boldsymbol{p}^i_{jkm}\}$ where $\boldsymbol{p}^i_{jkm}$ is the m-th path vector (node-sequence) between border nodes v^i_j and v^i_k in domain i, $1 \leqslant m \leqslant K$. Here only paths which satisfy wavelength continuity constraint are considered in this step, i.e., at least one common wavelength has to be available along all links of a given path. This can be done by modifying the generic Dijkstra algorithm to track the

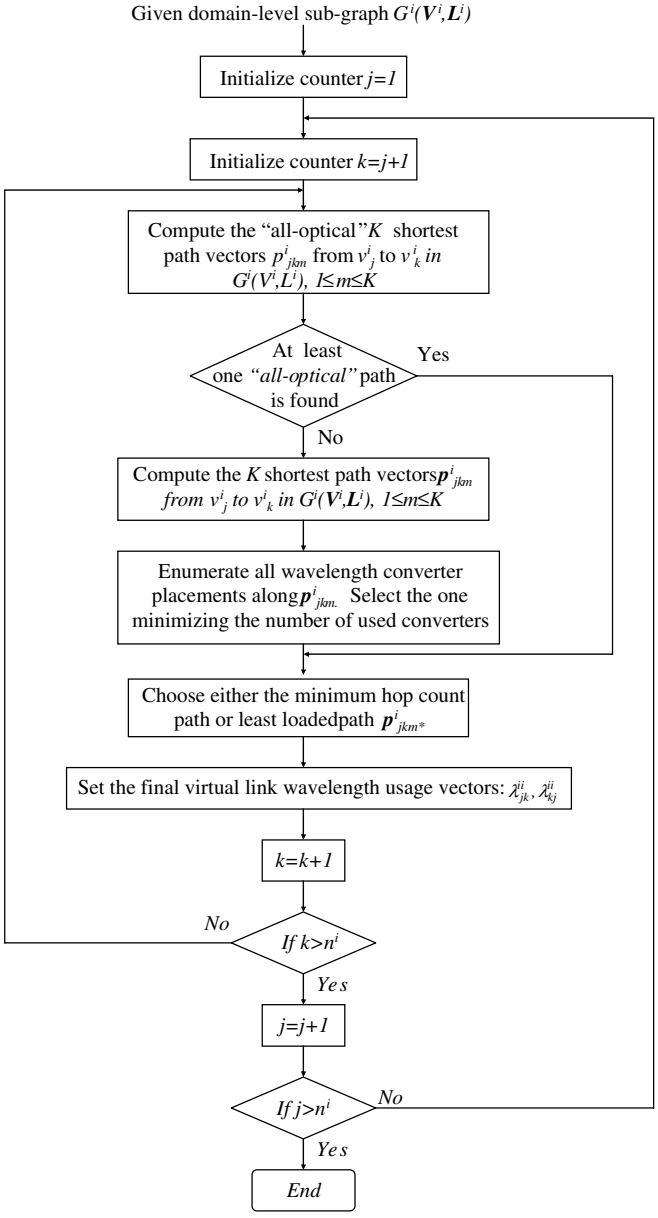

Fig. 2. Modified mesh topology abstraction algorithm (with full wavelength conversion).

wavelength availability vector, i.e., a newly found node is the closest node that also satisfies wavelength continuity constraint. This is denoted as "*all-optical*" *K*-shortest path (Fig. 2). Note if no such a path can be found, then *opto-electronic* paths must be computed as follows. Namely the *K*-shortest path algorithm is first run to generate a set of candidate paths between the border node pairs. All converter combinations along these paths are then searched to minimize the number of used converters. Finally, either the minimum hop count path or least loaded path is selected and the resultant wavelength availability vector is computed for the virtual link between the border nodes, λ^{ii}_{jk}, $1 \leqslant j, k \leqslant b^i$, as shown in Fig. 3.

Overall full-mesh abstraction provides more accurate intra-domain usage state versus the simple node scheme, albeit at the cost of significant computational complexities (RAL node) and higher inter-domain routing loads. Namely, inter-domain nodes must maintain state for $O(n^i(n^i-1)) = O((n^i)^2)$ virtual links for domain *i*, in addition to the actual physical inter-domain links. This yields a total storage/update complexity of $O(|\boldsymbol{E}| + \sum^i (n^i)^2)$ across all domains, a notable increase over the simple node scheme.

3.3. Hierarchical routing and triggering update policies

As stated earlier, the proposed routing scheme uses a two-level hierarchical *link-state* approach for multi-domain DWDM lightpath RWA. The first level runs the modified GMPLS OSPF-TE protocol between OXC nodes to maintain full wavelength-level state for all links, i.e., extensions for DWDM links [2]. Here *link-state advertisement* (LSA) updates are generated via a *significance change factor* (SCF) triggering policy, as studied for (single-domain) IP routing networks [13] and DWDM networks [19]. In particular, the results in [19] have shown the overall benefit of this approach versus periodic update triggering, e.g., lower routing update messaging loads with comparable or better blocking probability. In the SCF-based scheme, DWDM LSA updates are flooded to all neighboring nodes if the *relative* change in free wavelengths on a node's link exceeds the SCF value and the duration since the last update exceeds a *hold-down timer* (HT). These LSA updates contain full wavelength vectors indicating the free/reserved wavelengths on the link. Sequence numbers are also used to prevent infinite forwarding, i.e., "clamp-down" effect. Meanwhile the second level of *link-state* routing operates between border nodes and exchanges inter-domain state. This can be implemented, for example, by using two-level OSPF-TE link-state routing to provide a global summary of link resources, including physical and virtual links.

Carefully note that at the inter-domain level, additional triggering considerations are also needed for virtual link updates. In essence here it is not effective to strictly deploy either an SCF or timer-based update triggering policy as the inherent link quantities here are computed and not physical. For example, if pursuing a strictly SCF-based mechanism it is clearly unfeasible to compute domain-level abstractions every time there is a change in internal link state. Conversely using a fixed timer-based policy may help lower compute complexities but resultant routing loads may be significantly higher. As a result the proposed solution herein (as originally developed in [17]) combines the above two strategies. Namely, the domain RAL nodes are required to periodically compute domain abstractions at the expiry of an *inter-domain hold-down timer* (IHT). Subsequently, virtual link LSA updates are only propagated if there is sufficient relative change in wavelength state on the computed abstract links.

Also note that for the opto-electronic case, additional converter state can also be propagated, i.e., to acquire a

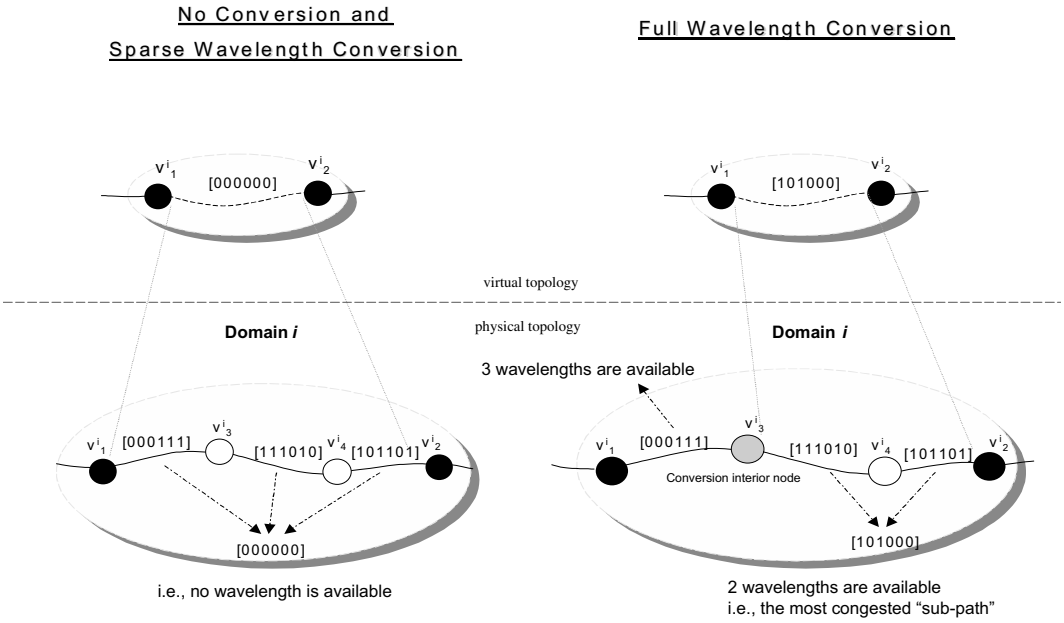

Fig. 3. Full-mesh abstraction in the presence of full wavelength conversion.

synchronized node converter state. This is achieved by adding a new converter-count field to the LSA update for physical intra-domain/inter-domain links. Namely, at the first routing level (intra-domain), an interior OXC node generating LSA updates can simply insert its latest per-node wavelength counts in this field. Hence the available converter count at an interior node can be determined by simply tracking the value of this field in the latest LSA updates from this node. Meanwhile at the second routing level (inter-domain), converter count fields can also be attached to LSA updates *for physical* inter-domain links, i.e., representing the latest available converter count at the sourcing border node. Note that this field is not used on inter-domain virtual links since the abstraction of converter state inside a domain is represented by the wavelength availability vector of virtual links, as per Section 3.1.2.

4. Inter-domain lightpath provisioning

Distributed lightpath provisioning performs inter-domain route selection/wavelength assignment and setup – a very challenging problem given the limitations of partial domain state. Note that this has been studied in [7] for all-optical and sparse wavelength conversion (opto-electronic) networks. The more challenging case of full wavelength conversion is treated here.

In order to maximize effectiveness, the proposed scheme leverages both inter-domain path computation and distributed RSVP-TE *loose route* (LR) signaling [2]. Again, RSVP-TE is a modified standardization of RSVP signaling specifically designed to reserve resources on circuit-switching links (DWDM, SONET/SDH). Namely, consider the setup sequence for an inter-domain lightpath originating at an *interior* source OXC. This node first sends a query message to its nearest border node to compute a LR domain sequence to the destination OXC. The queried border node then returns a LR sequence specifying the *end-to-end* border OXC node sequence to the destination domain, e.g., egress border OXC at the source domain, all ingress/egress border OXC nodes at intermediate domains, and final ingress border OXC at the destination domain. Note that if the source node is also a border node, it can perform LR computation locally.

Next the computed LR is packaged into a RSVP-TE *PATH* signaling message at the source node and sent downstream to resolve the end-to-end path. Here intermediate ingress border nodes perform *explicit route* (ER) expansion (and possibly wavelength selection) on the incoming LR sequence to resolve the set of exact nodes/links to the intermediate egress border OXC nodes in their domain. Note that the destination domain ingress border OXC node performs ER expansion to the destination OXC. Specific details for the case of transparent and translucent lightpath provisioning are not presented here, and interested readers are referred to [7].

For the focus herein, the more general case of wavelength conversion inside a domain is notably more involved. First lightpath RWA must expand all route links *and* identify the specific wavelengths on each of these links (i.e., sub-paths). However, unlike translucent lightpath setup (i.e., with sparse conversion [7]), wavelength conversion may also be done at interior OXC nodes. Therefore the exact interior node locations performing conversion must be identified via two rounds of exhaustive search. Namely, conversion-enabled border OXC nodes performing conversion are first identified during LR expansion using the inter-domain database. Conversely, interior

OXC nodes performing conversion are computed thereafter during the ER expansion phase.

Consider the details for the inter-domain LR computation, first the *K*-shortest path algorithm is run between source and destination domain border nodes, i.e., over the virtual inter-domain topology graphs, i.e., $\boldsymbol{H_{sn}}(U_{sn}, E)$ or $H_{\text{mesh}}(\boldsymbol{U}, \boldsymbol{E} + \sum_i \{\boldsymbol{E}^i_{\text{mesh}}\})$. Next a candidate LR sequence is chosen using either one of two metrics – *minimum hop count* or *minimum converter count* [7]. Namely, the minimum hop count chooses the shortest end-to-end path (from the candidate *K*-shortest paths). Subsequently, all converter combinations along this LR are searched to minimize the number of used converters. This step identifies the exact border node locations for wavelength conversion, i.e., only selected border nodes with wavelength conversion are required to actually use them. Overall this scheme exploits wavelength converters to setup shorter inter-domain paths. Conversely, the minimum converter count chooses the end-to-end path with the minimum number of border node converters (i.e., based upon available *inter-domain* routing database). Herein the *minimum hop count* metric is adopted to save crucial inter-domain node and link resources. Note that future efforts can study more advanced link cost metrics, e.g., adaptations of those proposed in [14].

Finally, consider setup signaling implications here, Fig. 4. Upon LR selection, the computed node sequence is inserted into a downstream RSVP-TE *PATH* signaling message. Again this message contains an initialized "all-ones" path availability vector *and* explicitly identifies all border nodes which must perform wavelength conversion. Hence when an ingress border node receives a *PATH* message it must perform ER expansion and identify converter locations at all interior nodes, i.e., using the local domain's physical topology/resource database, i.e., $G^i(\boldsymbol{V}^i, \boldsymbol{L}^i)$. Similar to the LR expansion the minimum hop count method is also adopted here. Moreover, since intra-domain links have shorter span distances and lower congestion degrees, ER expansion may pursue two possible strategies, as discussed further.

4.1. Strategy I: "all-optical" first

Here the ingress border node first computes an "all-optical" shortest path using modified Dijkstra algorithm, i.e., at least one common wavelength must be available on the route from ingress to egress border node. Note that this could possibly yield a longer path than the actual minimum hop count path. If no such a path is found, the *K-shortest* paths are computed and all associated converter combinations are searched (similar to LR expansion) to find feasible paths. Here a candidate path is chosen again using the *minimum hop count* metric. Overall this approach can help balance carried load across network links. For example, for domain *i* in Fig. 5, path v^i_1-v^i_5-v^i_6-v^i_2 is a possible candidate of this strategy and yields a wavelength availability vector of [101110].

4.2. Strategy II: shortest path first

This strategy tries to leverage converters directly to find a shorter path between an ingress/egress border node pair, i.e., to minimize wavelength usage. Namely, instead of searching for any feasible all-optical path first, the *K-shortest* paths and conversion locations are computed by using the same method as Strategy I. For example, in the domain shown in Fig. 5, this strategy will select path v^i_1-v^i_4-v^i_2 with conversion at v^i_4. It is clear that the resulting path is shorter here, i.e., less wavelengths used.

Now consider a conversion node receiving a *PATH* message. Necessarily this node must select a wavelength for all previous links in the expanded *PATH* LR sequence *up to* the last conversion OXC node (or source node). This can be done using either *most-used* (MU), *least-used* (LU), or *first-fit* (FF) wavelength selection, as per [7]. For example, node b_y in Fig. 4 selects wavelengths for all prior path links. Next, this conversion node must re-set the path availability vector to "all-ones" and check for a non-zero nodal wavelength converter count, i.e., $C > 0$. The *PATH* message propagated downstream only if a converter is available. This step effectively "re-generates" a lightpath at a conver-

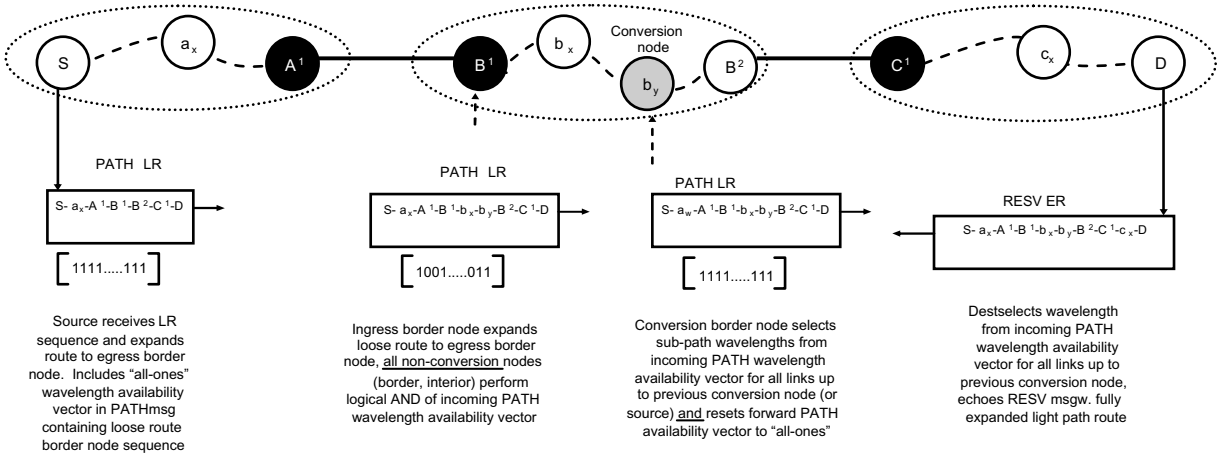

Fig. 4. Inter-domain lightpath signaling.

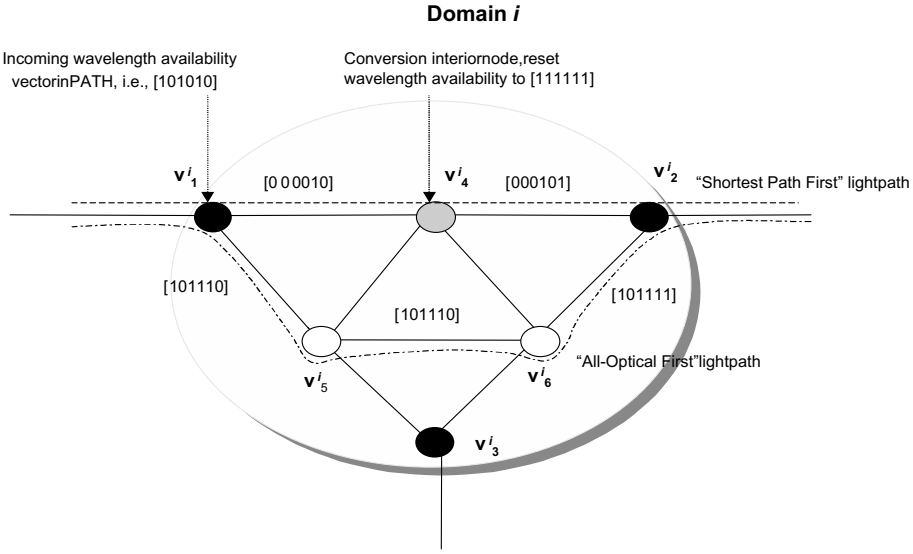

Fig. 5. Inter-domain ER expansion strategies.

sion node. Finally, the receiving destination OXC performs wavelength selection up to the last conversion node in the expanded *PATH* sequence, e.g., node *D* in Fig. 4 selects wavelengths for all links up to node b_y.

5. Performance evaluation

The multi-domain DWDM provisioning framework is tested using the *OPNET Modeler*™ discrete event simulation tool. Here an extensive inter-domain GMPLS protocols suite is developed and a 9-domain topology with 19 pairs of uni-directional inter-domain links is tested, see Fig. 6, as originally developed and tested in [7]. This network has good inter-domain connectivity, averaging 4.22 inter-domain links per domain (to stress inter-domain performance). Each link is assigned either $W = 8$ or 16 wavelengths and each node is assigned $C = 10$ converters (opto-electronic cases only). As for wavelength assignment schemes, MU is adopted for all the simulation runs as some

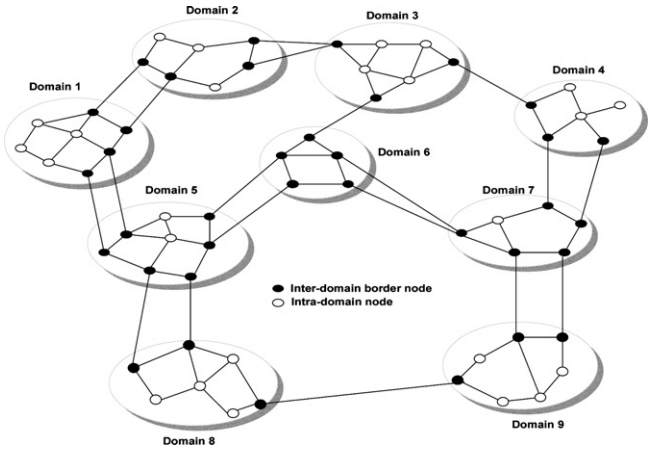

Fig. 6. Nine-domain test network topology.

initial results in [6] indicate that MU outperforms LU in multi-domain DWDM networks. Meanwhile for ER expansion strategies, Strategy I is used unless otherwise noted. All connections are generated between randomly selected domains using a 70–30% intra/inter-domain ratio. This is chosen to reflect practical settings which will likely field more intra-domain requests than inter-domain requests. Within a given domain the respective source and destination OXC nodes are also chosen randomly via a uniform distribution. Furthermore intra/inter-domain routing hold-down timers (*HT*,*IHT*) are set to 10 sec and the SCF is set to 10%. All runs are averaged over 200,000 connections and mean holding times are set to 600 s (exponential). Request inter-arrival times are also exponential and vary with loadings.

Inter-domain lightpath blocking results are first presented for $W = 8$ wavelengths in Fig. 7. Specifically, the two topology abstractions schemes are tested under no conversion (i.e., "all-optical"), sparse wavelength conversion, and full wavelength conversion. As expected, a graded relationship is observed with full wavelength conversion yielding the lowest blocking (for both full-mesh and simple node abstractions), followed by sparse wavelength conversion and no conversion. Here it is clear that a very moderate amount of conversion at the border nodes (i.e., sparse conversion) can yield sufficient gains, on par with full wavelength conversion at all nodes. This concurs with previous findings for wavelength conversion in single-domain networks [15]. More importantly, full-mesh abstraction gives a significant reduction in blocking probability versus simple node for all three scenarios. Similar results are also seen for $W = 16$ wavelengths, shown in Fig. 8, e.g., at 91 Erlang load the full-mesh scheme gives about 55% lower inter-domain blocking with full conversion. Blocking reduction for $W = 8$ wavelengths is also good, averaging 45–55%. Note that full-mesh abstraction

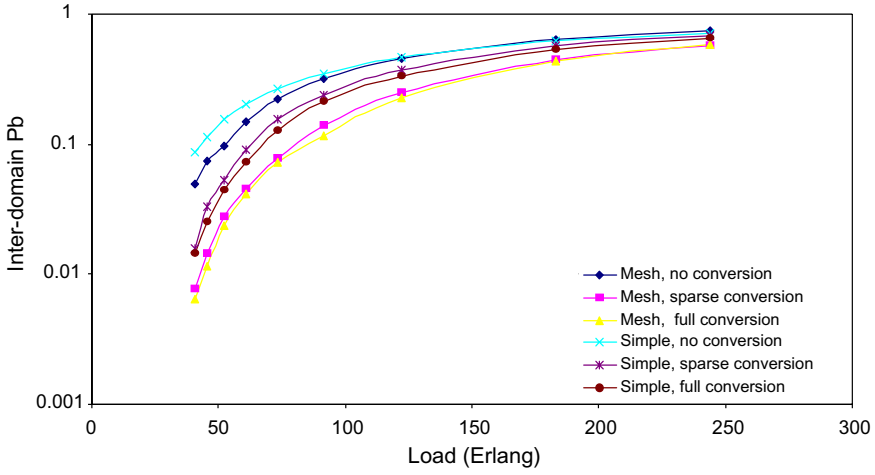

Fig. 7. Inter-domain lightpath blocking ($W = 8, C = 10$).

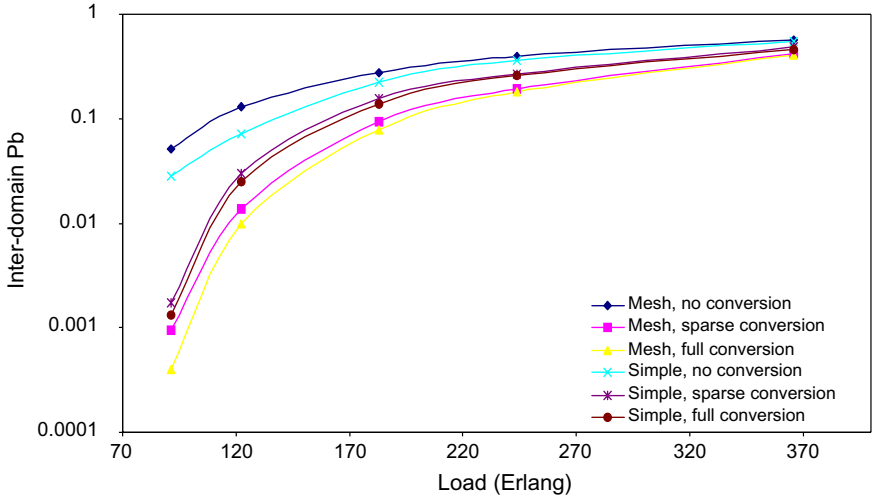

Fig. 8. Inter-domain lightpath blocking ($W = 16, C = 10$).

with sparse conversion also outperforms simple node with full conversion. This indicates that advanced software-based routing in conjunction with sparse conversion can preclude costly full conversion in operational networks. It is also noted that *intra-domain* lightpath blocking rates are very low and are not shown here. Moreover, additional runs with larger inter-domain topologies (more intra-domain nodes) reveal similar findings.

Subsequently, the blocking performances for the two ER expansion strategies (detailed in Section 3.3) are tested for $W = 16$ with full conversion, Fig. 9. Again the findings confirm that full-mesh topology abstraction has a lower blocking under both strategies. More importantly, the results show that Strategy I gives significantly lower blocking than Strategy II. This is due to the fact that a load balancing effect can be achieved via Strategy I whereas in Strategy II, traffic load tends to be restricted to shortest paths, thus having a much higher chance of being blocked.

Next, the fractional blocking of computation failures (i.e., the number of failed LR expansions) and signaling failures are plotted versus connection mean inter-arrival time, as shown in Fig. 10. Here, results show that more connections are blocked during the LR computation stage with full-mesh abstraction. In turn this yields much lower signaling blocking rates for RSVP-TE signaling and ER expansion with this particular scheme – a key advantage. To further gauge this particular saliency, the aggregate RSVP-TE error messages exchanged for $W = 8$ and $W = 16$ are shown in Fig. 11, including both *PATH ERR* and *RESV ERR* messages. Note that for the case of *RESV-ERR*, all downstream nodes from the error sourcing node (all the way to destination node) have to free resources (both wavelengths and converters) that have been earlier reserved by the failed connection. Clearly, this might trigger a sequence of *FREE* messages. Conversely, for the case of *PATH ERR*, only the source node has to be notified to drop the failed connection, see [2] for details on *RSVP-TE* signaling. From Fig. 11, it is clear that simple node abstraction ($W = 16$) generates nearly two times as many error messages, implying much higher control plane bandwidth

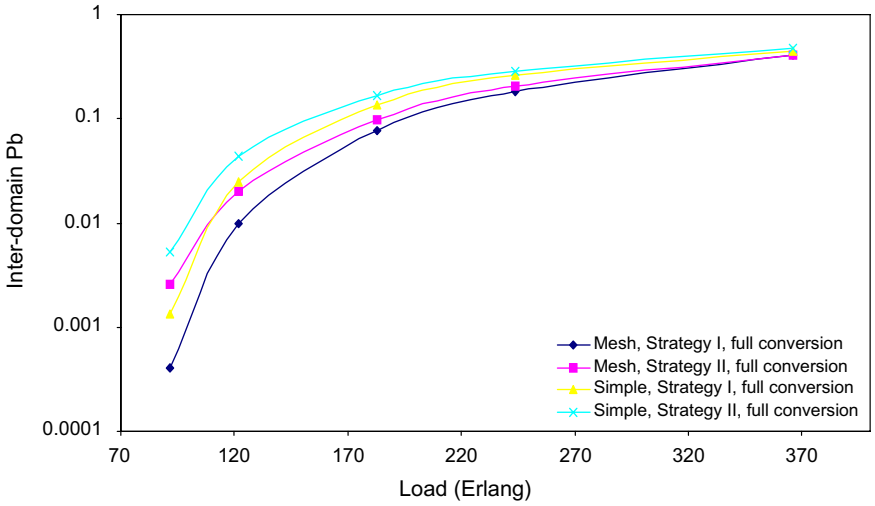

Fig. 9. Explicit route expansion strategies ($W = 16, C = 10$).

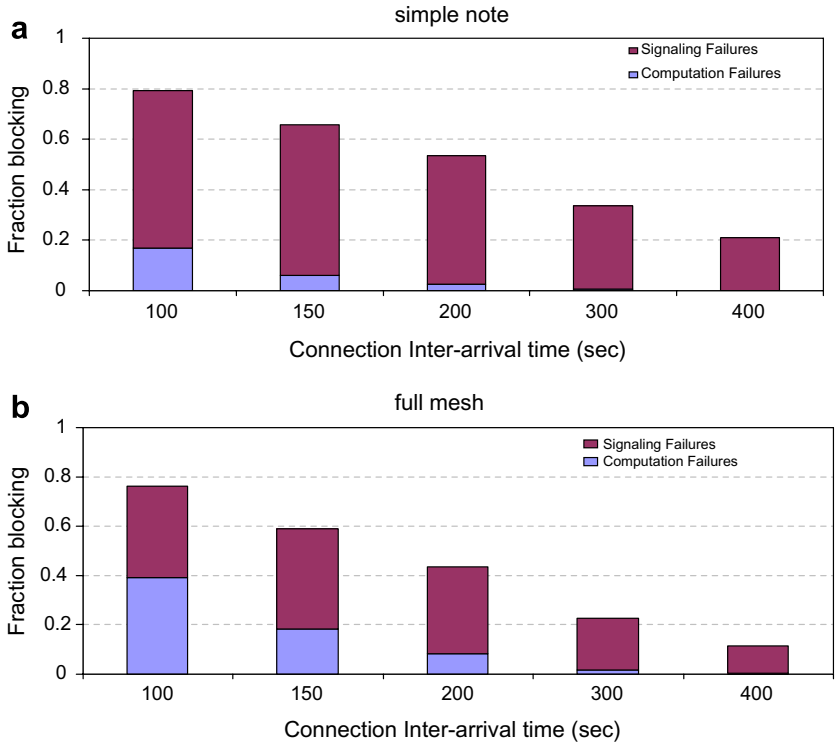

Fig. 10. Fraction blocking of signaling failures and computation failures ($W = 8$).

and CPU processing overheads. This is a key concern as control plane bandwidth is generally more limited and well below gigabit wavelength capacity, see [1].

In addition, the average number converters used per successful connection setup are plotted in Fig. 12, for both full-mesh and simple node abstraction. Here given the same amount of converters, full mesh abstraction is shown to yield much better converter resource efficiency. For example, at 183 Erlang load, the number of converters used by full mesh abstraction averages about 0.87 converter/successful connection. This value is almost 7 times higher than that with simple node abstraction, which averages only about 0.10. These results also provide insights into why better blocking performance can be achieved via full mesh abstraction, i.e., better utilizing converters.

The inter-domain OSPF-TE routing loads (measured in inter-domain LSA/s) are also shown in Fig. 13, for the full-mesh and simple node schemes. These findings confirm a significantly higher routing load for the full-mesh abstraction scheme, almost four times higher. This is expected as there are 38 physical inter-domain links (i.e., 19 pairs of uni-directional links) and 138 virtual links (as per mesh abstraction for all domains in Fig. 4), approximately four times larger. In order to address this concern and achieve

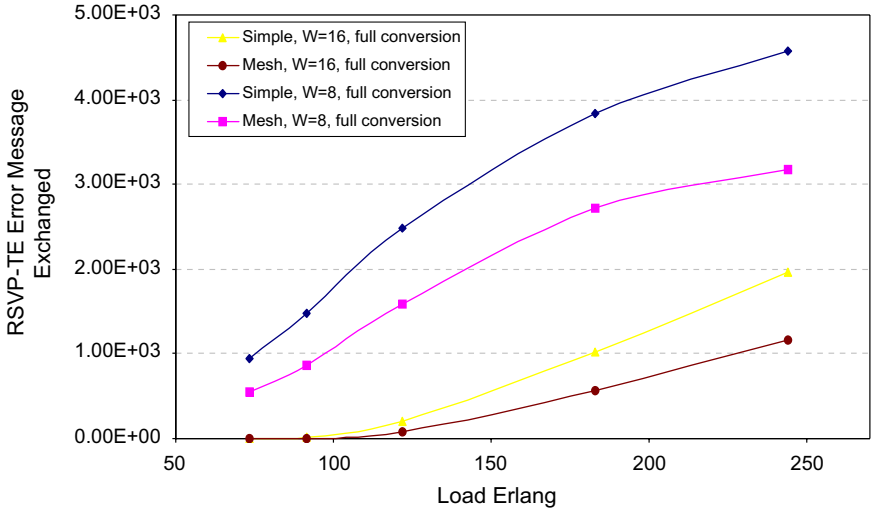

Fig. 11. RSVP-TE error messages exchanged ($C = 10$).

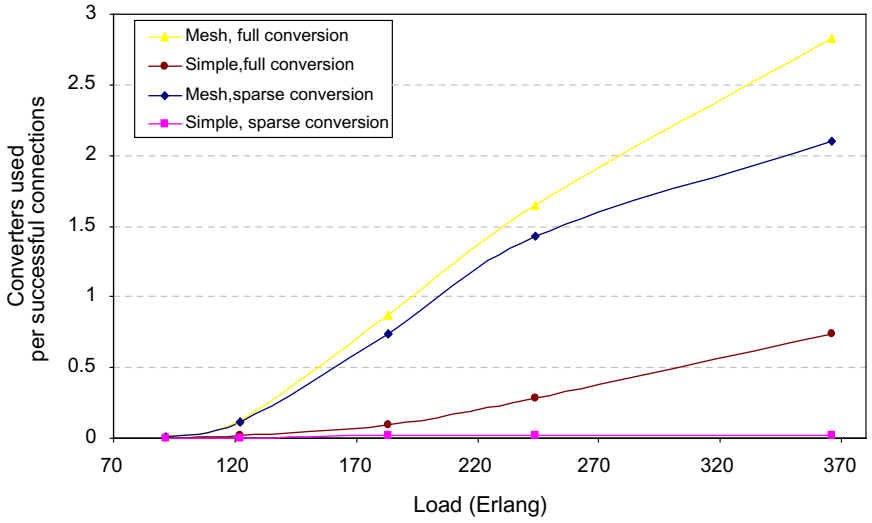

Fig. 12. Converter usage ($W = 16, C = 10$).

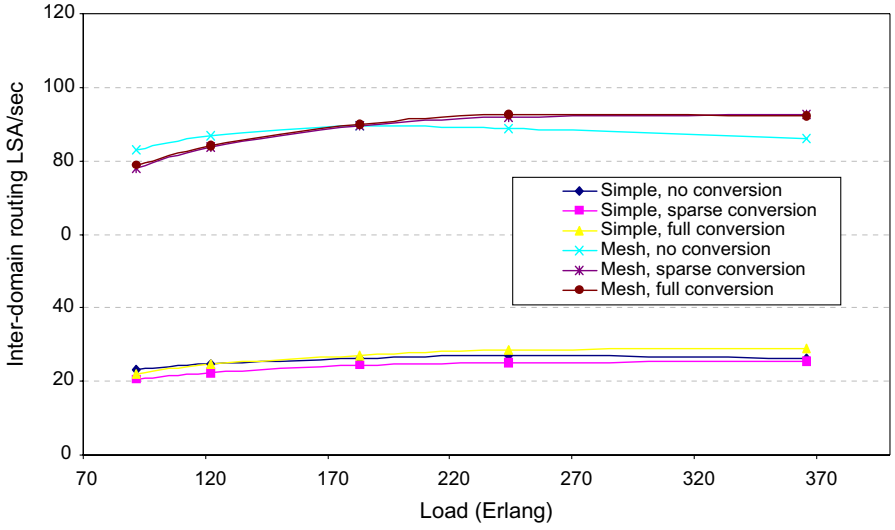

Fig. 13. Inter-domain routing LSA load ($W = 16$).

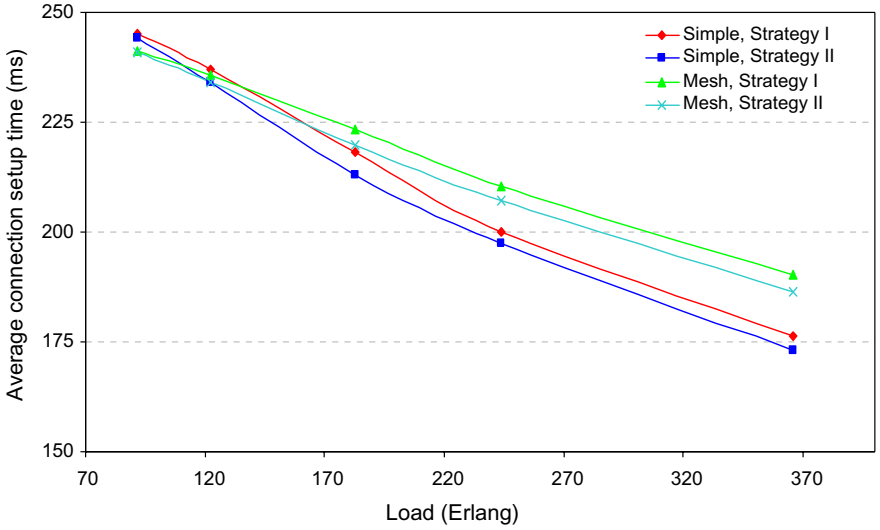

Fig. 14. Connection setup latencies for full wavelength conversion ($W = 16, C = 10$).

a better tradeoff between inter-domain routing load and blocking performance, more efficient star abstractions can be considered, and this will be the focus of future studies. Also note that there is minimal difference in routing load between the all-optical and opto-electronic schemes, and that the IHT time quickly limits routing overheads at increased connection loadings, i.e., "clamp-down" effect.

Finally, lightpath connection setup times are also a key concern for carriers operating across multiple domains. Here, Fig. 14 plots the setup times for full wavelength conversion for different connection loadings and $W = 16$ wavelengths and $C = 10$ converters/node. The switch setup times are set to 20 ms, a reasonable value for OXC designs based upon *micro-electro-mechanical systems* (MEMS) technology. As seen from these results the overall setup delay decreases with load for all schemes. This is expected as connection blocking increases with loading, particularly along longer attempted lightpath routes. More importantly these findings also show that full-mesh abstraction gives higher setup delays at lower (carrier operation regime) loading, about 8% higher. In addition Strategy I also yields higher setup delays versus Strategy II. By and large these results are expected for several reasons. Foremost, mesh abstraction provides more accurate global network state, which benefits load-balancing type routing methodologies which tend to give longer routes at lower loadings. Moreover, Strategy I directly pursues a load-balancing approach versus the shortest path approach of Strategy II.

6. Conclusions and future directions

This paper presents a comprehensive framework for inter-domain lightpath provisioning in DWDM networks with varying degrees of wavelength conversion, i.e., no conversion, sparse conversion and full conversion. A hierarchical routing setup is presented along with two topology abstraction schemes, simple node and full-mesh. Detailed inter-domain routing/triggering policies and RWA and signaling procedures are also defined. The results show a notable reduction in inter-domain lightpath blocking with full-mesh abstraction. In addition, commensurate RSVP-TE setup signaling loads are also much lower here and improved wavelength converter efficacy is also observed. However, all of these gains require inter-domain routing overheads. Finally, the results also show that networks with sparse conversion when augmented with advanced routing/topology abstraction algorithms can achieve better blocking performance than those with full conversion only (i.e., simple node). Future efforts will look at advanced star abstractions along with inter-domain crankback signaling strategies. The more challenging, yet realistic, concern of multi-layer SONET/DWDM domain interfacing will also be addressed as well.

Acknowledgements

This research has been supported in part by the US *Department of Energy* (DOE) Office of Science and the United States *National Science Foundation* (NSF) *Computer and Information Sciences Division* (CISE) under Grant No. CNS-0448225. The authors are very grateful to these agencies for their generous support.

References

[1] N. Ghani. J. Pan, X. Cheng, Metropolitan optical networks, Optical Fiber Telecommunications IV, Academic Press, London, March 2002, pp. 329–403 (Chapter 8).

[2] G. Bernstein, B. Rajagopalan, D. Saha, Optical Network Control-Architecture, Protocols and Standards, Addison Wesley, Boston, 2003.

[3] H. Zang, J. Jue, B. Mukherjee, A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks, Optical Networks Magazine 1 (1) (2000) 47–60.
[4] W. Alanqar, A. Jukan, Extending end-to-end optical service provisioning and restoration in carrier networks: opportunities, issues, and challenges, IEEE Communications Magazine 42 (1) (2004) 52–60.
[5] I. Iliadis, Optimal PNNI complex node representations for restrictive costs and minimal path computation time, IEEE/ACM Transactions on Networking 8 (4) (2000) 493–506.
[6] X. Bruin, et al., Hierarchical routing with QoS constraints in optical transport networks, in: 3rd IFIP-TC6 Networking Conference, Athens, Greece, May 2004.
[7] Q. Liu, M.A. Kök, N. Ghani, A. Gumaste, Hierarchical routing in multi-domain optical networks, Computer Communications 30 (1) (2006) 122–131.
[8] J. Vasseur et al., Inter-area and inter-AS MPLS traffic engineering, IETF Draft draft-vasseur-ccamp-inter-area-as-te-00.txt (2004).
[9] B. St. Arnaud et al., BGP optical switches and lightpath route arbiter, Optical Networks Magazine 2 (2) (2001) 73–81.
[10] X. Yang, B. Ramamurthy, Inter-domain dynamic routing in multi-layer optical transport networks, in: IEEE GLOBECOM 2003, San Francisco, CA, December 2003.
[11] S. Sanchez-Lopez, et al., A hierarchical routing approach for GMPLS-based control plane for ASON, in: IEEE ICC 2005, Seoul, Korea, June 2005.
[12] P. Ho, H.T. Mouftah, A novel survivable routing algorithm for segment shared protection in WDM networks with partial wavelength conversion, IEEE Journal of Selected Areas on Communications 22 (8) (2004) 1539–1548.
[13] R. Alnuweri et al., Performance of new link state advertisement mechanisms in routing protocols with traffic engineering extensions, IEEE Communications Magazine 42 (5) (2004) 151–162.
[14] Q. Ma, P. Steenkiste, On path selection for traffic with bandwidth guarantees, in: International Conference on Network Protocols (ICNP'97), Atlanta, GA, October 1997.
[15] C. Assi, A. Shami, Z. Zhang, M. Ali, Impact of wavelength converters on the performance of optical networks, SPIE Optical Network 3 (2) (2002) 16–23.
[16] M.Yannuzzi, et al., Interdomain RWA based on stochastic estimation methods and adaptive filtering for optical networks, in: IEEE GLOBECOM 2006, San Francisco, November 2006.
[17] Q. Liu, et al., Application of topology abstraction techniques in multi-domain optical networks, in: IEEE ICCCN 2006, Arlington, VA, October 2006.
[18] G. Liu, C. Ji, V. Chan, On the scalability of network management information for inter-domain light-path assessment, IEEE/ACM Transactions on Networking 13 (1) (2005) 160–172.
[19] A. Shami et al., Connection management protocols for dynamic lightpath provisioning in future WDM networks, Photonic Network Communications Journal 6 (1) (2003) 25–32.

Qing Liu received the B.E. degree in telecom engineering and M.E. degree in computer science from the Nanjing University of Posts and Telecommunications, China, in 2001 and 2004, respectively. He is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of New Mexico. Before he joined UNM, he had worked with Motorola China as a summer intern and Lucent Technologies China as Software Engineer. His research interests include optical networks, multimedia communication, and wireless networks.

Dr. Nasir Ghani has gained a wide range of experience in the telecommunications area and has held senior positions at Nokia, IBM, Motorola, Sorrento Networks, and Tennessee Tech University. Currently he is an Associate Professor in the Department of Electrical and Computer Engineering at the University of New Mexico, where he is actively involved in a wide range of funded research projects in the area of optical networks and cyber-infrastructures. Dr. Ghani has published over 80 journal and conference papers, several book chapters, various standardization proposals, and has two patents granted. He recently served as a co-chair for the optical networking symposia for *IEEE ICC 2006* and *IEEE GLOBECOM 2006* and is a program committee member for *OFC 2007* and *OFC 2008*. Furthermore he has been a program committee member for numerous IEEE, SPIE, ACM, and IEC conferences and has served regularly on NSF, DOE, and international panels. He is an associate editor of the *IEEE Communications Letters* journal and has guest-edited special issues of *IEEE Network*, *IEEE Communications Magazine*, and *Cluster Computing*. Dr. Ghani is a recipient of the prestigious NSF CAREER Award and is a Senior Member of the IEEE. He received the Ph.D. degree in Electrical and Computer Engineering from the University of Waterloo, Canada.

Dr. Nageswara S. V. Rao is currently UT-Battelle Corporate Fellow at Oak Ridge National Laboratory where he joined in 1993. He received ME from the School of Automation, Indian Institute of Science, Bangalore in 1984, and the Ph.D. in computer science from Louisiana State University in 1988. He was an assistant professor of computer science at Old Dominion University during 1988–1993. His research interests include network transport dynamics, statistical approaches to transport control, information and sensor fusion, robot navigation, and fault diagnosis. He is an associate editor for Pattern Recognition Letters, International Journal of Distributed Sensor Networks, and Information Fusion.

Dr. Ashwin Gumaste is a faculty member in the Department of Computer Science and Engineering at the Indian Institute of Technology Bombay. He was previously with Fujitsu Laboratories (USA) Inc as a Member of Research Staff in the Photonics Networking Laboratory (2001–05). 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 forty pending U.S. and EU patents and has published close to sixty papers in referred conferences and journals. He has authored three books in broadband networks, namely *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* (John Wiley). Dr. Gumaste 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 *IEEE ICC, IEEE Globecom, IEEE Broadnets, IEEE ICCCN, Gridnets,* etc. Dr. Gumaste is also a guest editor for *IEEE Communications Magazine* and is the General Chair of the *1st International Symposium on Advanced Networks and Telecommunication Systems (ANTS 2007)* to be held in Bombay, India.

Manuel Lopez Garcia is currently working towards his Masters degree in the Department of Electrical and Computer Engineering at Tennessee Tech University. He received his Bachelors degree in Electrical Engineering from Tennessee Tech University in spring of 2006. His interests include networking and telecommunications, application design, and professional motorcycle racing.