A Novel Topology Aggregation Approach for Shared Protection in Multi-domain Networks

Dieu-Linh Truong

Departement of Data Communications and Computer Networks, School of Information and Communication Technology, Hanoi University of Technology, Hanoi, Vietnam

Brigitte Jaumard

Computer Science and Software Engineering, Concordia University, Montréal, QC, Canada

---

Abstract

Routing for shared protection in multi-domain networks is more difficult than that in single domain networks because of the scalability requirements. We propose a novel approach for shared protection routing in multi-domain networks where the key feature is a special Topology Aggregation. In this Topology Aggregation, only some potential intra-domain paths (intra-paths for short) are selected for carrying working and backup traffic between domain border nodes. The abstraction of each intra-path to a virtual edge makes the original multi-domain network to become an aggregated network. On the aggregated network, a single domain routing algorithms for shared protection can be applied for obtaining the complete routing solutions. The experiments show that the proposed approach is scalable. Moreover it is close to the optimal solution in single-domain networks and outperforms the previously proposed scalable solutions in multi-domain networks.

Key words: Topology Aggregation, Multi-domain networks, Protection.

Email addresses: `linhtd@soict.hut.edu.vn` (Dieu-Linh Truong), `bjaumard@cse.concordia.ca` (Brigitte Jaumard)

1. Introduction

Many studies have been published for connection protections against failures. Some of them propose protection models such as link, path, segment or p-cycle, the other concentrate on the problem of allocating working and backup resources. When dedicated protection is employed, the resource allocation task is simply finding diverse paths for working and backup connections and can be solved by different diverse path routing algorithms such as those in [1] and [2].

For the bandwidth saving purpose, shared protection has been proposed for link, path and segment protections [3] or even Overlapping Segment Protection [4], a segment protection model where working segments can overlap each other. In addition to the basic idea of link, segment, overlapping segment and path protection, shared protection for these models allows sharing bandwidth amongst backup elements. Backup elements can be backup link, segment or path, commonly referred hereafter as "backup segments". Working elements are working link, segment or path and are similarly called "working segments".

In order to guarantee 100% recovery of any single link or node failure, two backup paths/segments are allowed sharing bandwidth if and only if their working segments are link and node-disjoint. This condition is called *sharing condition*, see Fig. 1 for an illustration. In case (a), the working segment from v_1 to v_2, with requested bandwidth d_1, and the working segment from v_5 to v_6, with requested bandwidth d_2, are link and node disjoint. Their backup segments can share bandwidth over the common link (v_4, v_3) and the needed backup bandwidth on this link is $\max\{d_1, d_2\}$ in order to be able to protect both working paths. In case (b), the two working segments share node v_7, their backup segments cannot share backup bandwidth. The needed backup bandwidth on link (v_4, v_3) is $d_1 + d_2$, which is greater than in case (a). Hence, the amount of backup bandwidth to be reserved for a backup segment depends on the working segment to be protected as well as on the existing working and backup segments. This dependency makes the routing problem for shared protection complex.

Shared protection under static traffic has received a lot of interest. Several efficient solutions have been proposed, especially the well-known p-cycle initially introduced in [5] and further developed for segment protection in [6],[7]. However, network traffic todays changes dynamically, static traffic is no longer an appropriate assumption except for planning. For this reason,

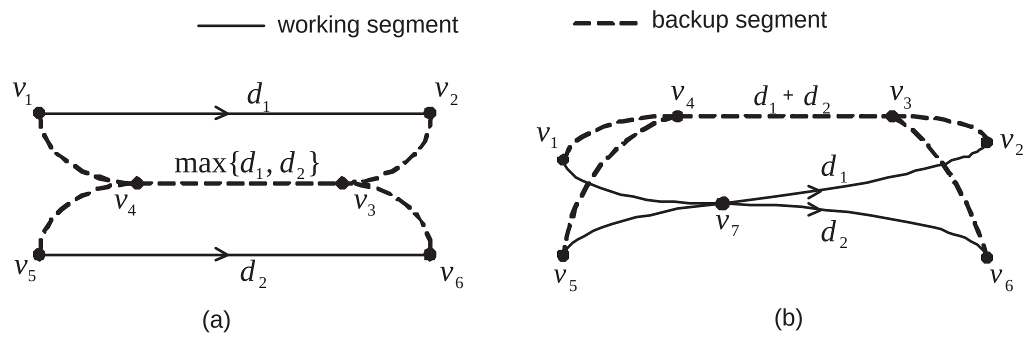

Figure 1: Examples of cases where two backup segments can share backup bandwidth (a) and cannot (b).

we focus only on dynamic traffic.

For a given new incoming request, the dynamic routing problem for shared protection consists of establishing a working path and associated backup segments for it, while minimizing the bandwidth they use. This routing should be done without any forecast on upcoming requests. Some optimal solutions for shared protection in single domain network has been proposed for example SCI model in [8] or the model in [4]. Several heuristics with smaller computational effort have also been proposed such as the works in [9],PDBWA and PIBWA in [10], SLSP-O in [11], CDR in [12], PROMISE in [13] or recursive shared segment protection in [14]. These works limit themselves in single domain networks because they need detailed information on bandwidth allocation on each network link for their complex bandwidth cost computations.

Shared protection for multi-domain networks is much more complex than that for single domain networks due to the network characteristics and size. A multi-domain network is made of the interconnection of several single-domain networks [15], see an illustration in Fig. 2a. In order to satisfy the *scalability requirements*, only the aggregated routing information can be exchanged amongst domains [16] by an Exterior Gateway Protocol such as BGP. Consequently, a given node is neither aware of the global multi-domain network topology nor of the detailed bandwidth allocation on each network link, although the complete routing information can still be available within each domain thanks to more frequent routing information updates performed by an Interior Gateway Protocol. This characteristic makes the current Shared protection routings for single domain networks are inapplicable for multi-domain networks.

Some works address the routing problem in multi-domain networks but very few solutions have been proposed for protection in multi-domain networks. These solutions have been analysed and evaluated in [17], [18]. Some of them, e.g. [19], [20], [21], do not take care of inter-domain link protection and turn the multi-domain protection into multiple intra-domain protections. The others tackle the scalability issue by using traditional Topology Aggregation approach such as nodal, full mesh or star model for aggregating each domain. The works in [22] and [23] proposed to use p-cycle protection at both intra-domain level and inter-domain level. Again multi-domain protection using p-cycle is a protection scheme for static or relatively stable traffic. Even in the stable traffic context, multi-domain p-cycle protection requires very high resource redundancy for protecting 100% links against failure. The works in [24], [25] proposed full mesh aggregations. Let denote a domain $N_m = (V_m, L_m)$, where V_m and L_m are the sets of nodes and links. In these researches, the domain is aggregated to become graph $G_m = (V_m^{\text{BORDER}}, V_m^{\text{2BORDER}})$ composed of a border node set V_m^{BORDER} and a virtual link set V_m^{2BORDER} (see Fig. 2b). A virtual link connects two border nodes of a domain and represents the set of domain internal paths running between these border nodes. A such path is called an intra-path. The multi-domain network becomes a so called inter-domain network. Each virtual link is then associated with approximative working and backup costs. Single domain routing algorithms for shared protection are used in this inter-domain network for finding the working and backup segments which are paths of virtual and inter-domain links. Virtual links are then mapped back to intra-paths in order to get the full end-to-end paths. In this paper, this approach is referred by "Route-and-Map" and denoted by RaM.

Although RaM offers good routing results and scalability, we found that the approximation made in working and backup cost computation leads the inter-domain routing to a solution that is different to the real one obtained after intra-domain routing. In this paper, we propose to eliminate the approximation in RaM. The idea is that: between each pair of border nodes, only some best intra-paths are used for carrying traffic. These intra-paths are then exposed as links at inter-domain level. The routing will be performed only in this inter-domain level. This approach can be seen as if we perform the mapping of intra-paths to virtual links first then routing. It is so

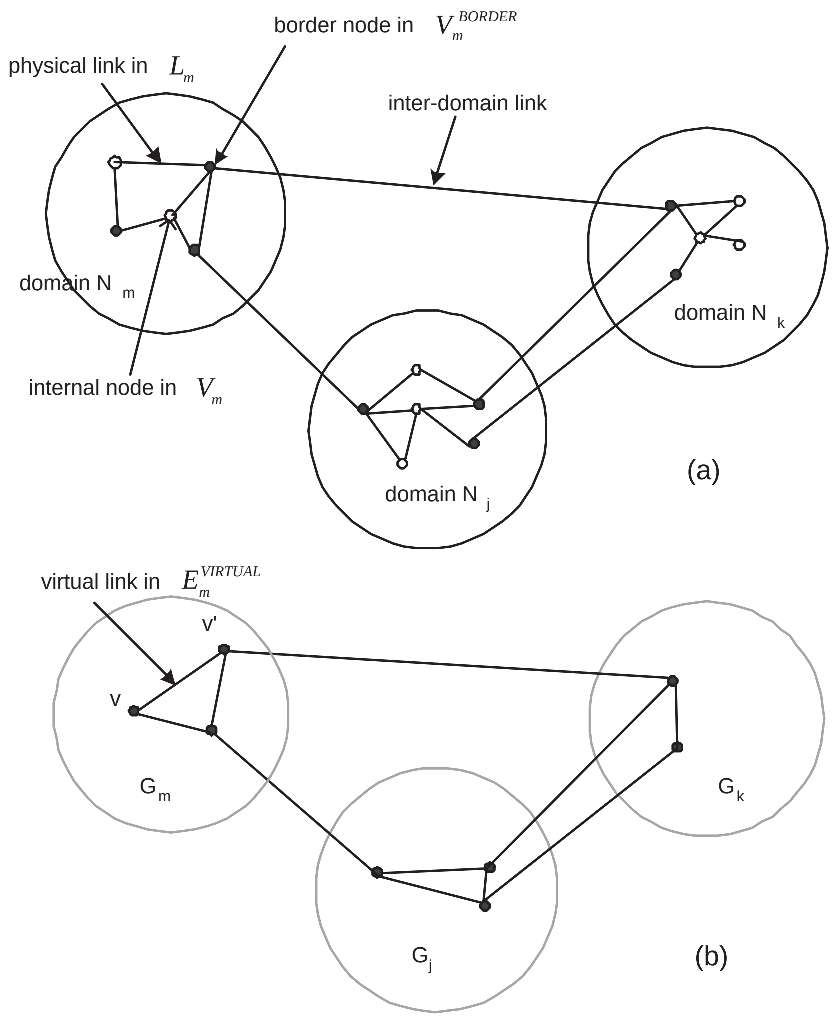

Figure 2: A multi-domain network (a) and its *inter-domain network* (b) obtained from Topology Aggregation.

called "Map-and-Route" or MaR for short. The advantage of this approach is that working and backup costs of intra-paths (i.e. links of inter-domain network) can be computed exactly and the routing is performed only once on the inter-domain network.

This paper is organized as follows. The next section provides general ideas of the proposed approach. Section 3 states the Mapping sub-problem in each domain, its exact and heuristic solutions as well as periodic Mapping refreshing. Section 4 describes the Routing solution and Section 5 discusses its scalability. The experimental results are shown and discussed in Section 6. Conclusions follow in Section 7.

2. General idea of Map and Route approach

Many intra-paths can carry traffic between a pair of border nodes. The main idea of MaR is that only some selected intra-paths will be used for this purpose. Those intra-paths are selected so that they support the best bandwidth saving and shared protection.

Let e represents a pair of border nodes of a domain network N_m, $\mathcal{P}_e^{\mathrm{W}}$ (resp. $\mathcal{P}_e^{\mathrm{B}}$) is the set of intra-paths that will be selected for carrying the working (resp. backup) traffic between the border nodes of e. The intra-paths in $\mathcal{P}_e^{\mathrm{W}}$ is called potential working intra-paths and those in $\mathcal{P}_e^{\mathrm{B}}$ are called potential backup intra-paths. We require that all selected intra-paths must be direct intra-path.

Definition 1. *A direct intra-path is a path that does not go through any intermediate border node other than the two end border nodes of the intra-path.*

Traffic crossing a domain can still go through an intermediate border node by taking more than one direct intra-path.

After being selected, each intra-path in $\mathcal{P}_e^{\mathrm{W}}$ and $\mathcal{P}_e^{\mathrm{B}}$ will be abstracted and handled as a single link called "virtual edge" (see Fig. 3, where border nodes are black filled nodes). Since a virtual edge is in one-to-one correspondence with an intra-path, we use the two terms alternatively depending on whether we are dealing with the abstracted (mapped) or detailed (intra-domain) level.

Although different virtual edges may share a common physical link, they manage separately their working and backup capacities. A virtual edge has its own working and backup capacities (see Fig. 4 for an illustration). The

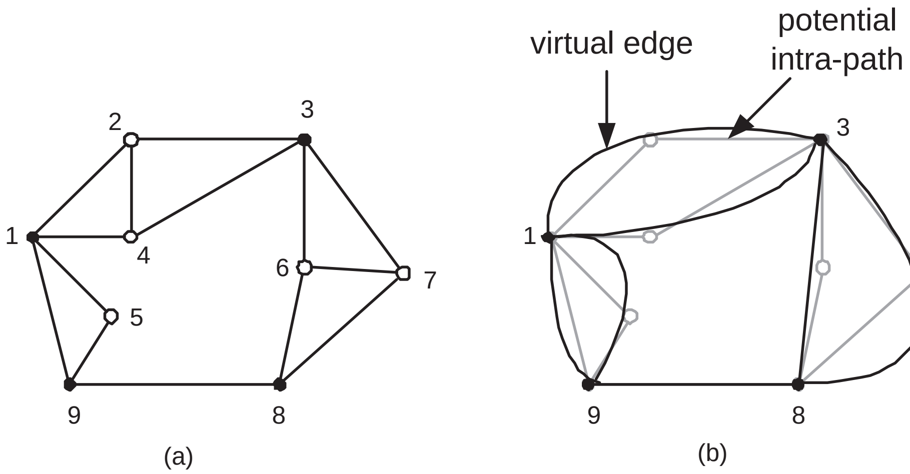

Figure 3: (a) An original domain; (b) the mapped domain with a maximum of 2 intra-paths/ pair of border nodes for both working and backup traffic.

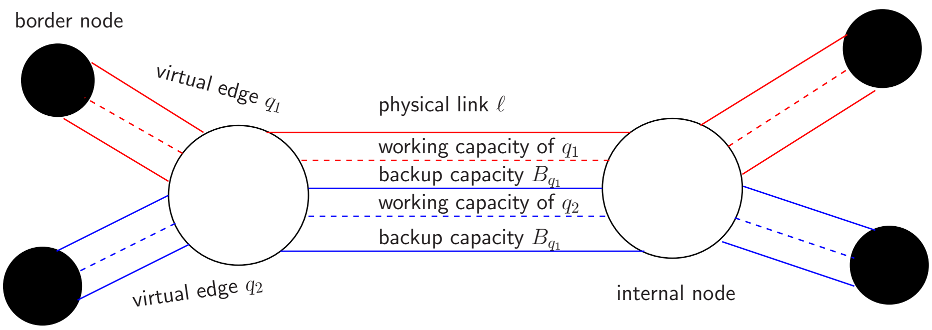

Figure 4: Example of two virtual edges that share physical link. Their B_{q_1}, B_{q_2} differs from the total backup capacity of link ℓ. The free capacities are not shown.

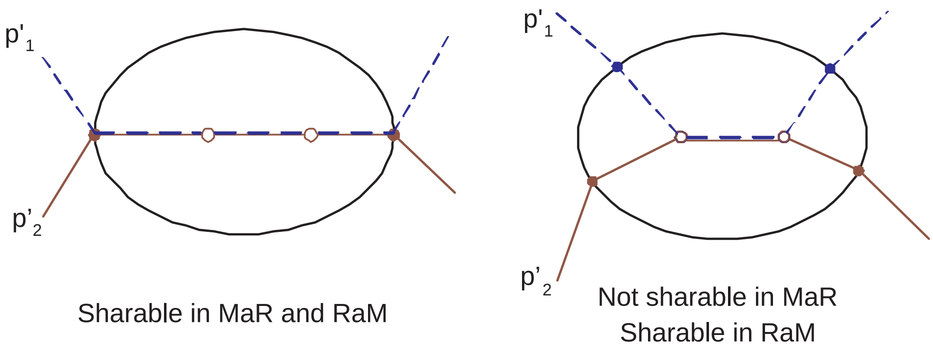

Figure 5: The cases where two backup segments p'_1, p'_2 can share and cannot share backup bandwidth under $\mathrm{M}a\mathrm{R}$ and $\mathrm{R}a\mathrm{M}$.

working capacity of a virtual edge is defined as the bandwidth occupied by the working paths routed along the entire intra-path associated with the virtual edge. Similarly, the backup capacity of a virtual edge is defined as the bandwidth occupied by the backup routes along the entire intra-path associated with the virtual edge. Therefore, the working (resp. backup) capacity of a physical link is the sum of, the working (resp. backup) capacities of all virtual edges that contain the link.

We also introduce the following new sharing rule, in which an intra-path is again handled as a single entity (see an illustration in Fig. 5).

> *Sharing rule*: Two backup segments are allowed to share bandwidth only if they go through an **identical intra-path**.

In other words, two backup segments either share bandwidth along their entire common intra-path or share no bandwidth. Backup segments over two intra-paths that differ by at least one link are not allowed to share bandwidth. Although some bandwidth sharing possibility is ignored on some particular physical links due to the introduced sharing rule, we will see later in Section 6 that it does not really decrease the sharing possibility. Moreover, we do not have to go down to the physical link level in order to identify some shareable bandwidth for protecting an intra-path, which would impair the scalability.

The routing can be performed in 2 phases:

- Mapping phase: Within each domain, two sets of potential working and backup intra-paths $\mathcal{P}_e^{\mathrm{W}}, \mathcal{P}_e^{\mathrm{B}}$ are selected for each pair of border nodes

e. Each intra-path is then abstracted as a single "virtual edge". The multi-domain network resulting from this abstraction is called "mapped network".

- Routing phase: The working and backup segments are computed in the mapped network by a single domain survivable routing algorithm.

The Mapping is performed once for a long term use and should be refreshed for updating the potential intra-paths only when the current ones are saturated. The Routing is performed uniquely in the whole mapped network and there is no need to go down to the intra-domain level for identifying the intra-paths within each domain as they are in one-to-one correspondence with virtual edges. Both phases will be discussed in detail in the next sections.

3. Mapping

3.1. Mathematical model

The Mapping consists of identifying a set of potential working intra-paths and a set of potential backup intra-paths between each pair of border nodes. Such a Mapping is performed independently in each domain. Let V_m^{2BORDER} be the set of pairs of border nodes of domain N_m, the Mapping problem for N_m is stated as follows.
Given:

- n^{W} and n^{B} the maximum numbers of potential working and backup intra-paths needed for each pair of border nodes;

- n_e the number of *direct* intra-paths between pair of border nodes e.

Let $n_e^{\text{W}} = \min\{n_e, n^{\text{W}}\}$ and $n_e^{\text{B}} = \min\{n_e, n^{\text{B}}\}$. They are the exact number of potential working and backup intra-paths to be selected for carrying traffic between a pair of border nodes e of N_m. We need to identify:

- $\mathcal{P}_e^{\text{W}} = \{q_{e,i}^{\text{W}}, i = 1..n_e^{\text{W}}\}$, the set of potential working intra-paths for e;

- $\mathcal{P}_e^{\text{B}} = \{q_{e,i}^{\text{B}}, i = 1..n_e^{\text{B}}\}$, the set of potential backup intra-paths for e.

Since the routing aims to minimize the total working and backup cost of each request, following selection criteria are defined for encouraging the intra-paths supporting this objective.

Criterion 1. *A potential working intra-path should have small working cost and maintain enough residual bandwidth for future connection allocation.*

This criterion aims to minimize working cost and also balance network load. Let assign to each physical link a weight which is the inverse of the residual capacity of the link. A potential intra-path should be the weighted shortest path. From the global viewpoint, the set of all potential intra-paths to be selected for a domain should thus minimize their total weighted length, which leads to:

$$\min \sum_{e \in V_m^{\text{2BORDER}}} \sum_{q \in \mathcal{P}_e^{\text{W}}} \sum_{\ell \in q} \frac{1}{c_\ell^{\text{res}}}. \tag{1}$$

Criterion 2. *A potential backup intra-path should have small backup cost and maintain enough residual bandwidth for future connection allocation.*

From a global sight, a backup segment uses an homogeneous amount of bandwidth along an intra-path. Hence, this criterion is interpreted similarly to Criterion 1, which leads to:

$$\min \sum_{e \in V_m^{\text{2BORDER}}} \sum_{q \in \mathcal{P}_e^{\text{B}}} \sum_{\ell \in q} \frac{1}{c_\ell^{\text{res}}}. \tag{2}$$

Criterion 3. *The potential working intra-paths should be selected so as to increase the possibility of finding pairwise disjoint working intra-paths.*

This criterion originates from the fact that backup segments can share bandwidth only if their working segments are disjoint, according to the segment sharing condition. The criterion is interpreted as maximizing the number of pairs of disjoint working intra-paths:

$$\max \sum_{\substack{q_1 \in \mathcal{P}_{e_1}^{\text{W}}, q_2 \in \mathcal{P}_{e_2}^{\text{W}}, \\ e_1, e_2 \in V_m^{\text{2BORDER}}}} \delta_{q_1}^{q_2}. \tag{3}$$

Criterion 4. *The possibility that a pair of border nodes is topologically protectable by another pair of border nodes should be maximized.*

The topological protect-ability between two pairs of border nodes is defined as follows:

Definition 2. *A pair of border nodes is said topologically protect-able by another pair of border nodes if there exists a potential working intra-path of the first pair that is link and node disjoint with a potential backup intra-path of the second pair.*

In the topology point of view, the potential working intra-path can be thus protected by the potential backup intra-path cited in the definition. Let us denote $\delta^{e}_{e'}$, the topological protect-ability of border node pair e' again the border node pair e, then:

$$\delta^{e}_{e'} = \begin{cases} 1 & \text{if } \exists q \in \mathcal{P}^{\mathrm{W}}_{e}, q' \in \mathcal{P}^{\mathrm{B}}_{e'} : q \cap q' = \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Criterion 4 is justified as follows. In the case of slightly inter-connected multi-domain networks, we may need to route a working segment of a request over a particular pair of border nodes and the corresponding backup segment over another pair of border nodes. These two pairs of border nodes should provide two intra-paths that are disjoint to each other, otherwise the considered backup segment would have some common links or nodes with its working segment and thus could not protect the working segment. From the global viewpoint, this criterion is interpreted as maximizing the number of sets of two pairs of border nodes that are topologically protect-able one another:

$$\max \sum_{e,e' \in V_m^{\text{2BORDER}}} \delta^{e}_{e'}. \tag{5}$$

When (5) gives more than one optimal solution, we break the ties by maximizing the total number of disjoint working and backup intra-paths, which is similar to Criterion 3:

$$\max \sum_{\substack{q \in \mathcal{P}^{\mathrm{W}}_{e}, q' \in \mathcal{P}^{\mathrm{B}}_{e'}, \\ e,e' \in V_m^{\text{2BORDER}}}} \delta^{q}_{q'}. \tag{6}$$

The Mapping is clearly a multi criteria optimization problem. In order to solve this problem, we combine de 4 criteria together:

$$\min\Bigg(\mu_1 \sum_{e\in V_m^{\text{2BORDER}}} \sum_{q\in\mathcal{P}_e^{\text{W}}} \sum_{\ell\in q} \frac{1}{c_\ell^{\text{res}}} + \mu_2 \sum_{e\in V_m^{\text{2BORDER}}} \sum_{q\in\mathcal{P}_e^{\text{B}}} \sum_{\ell\in q} \frac{1}{c_\ell^{\text{res}}} - \mu_3 \sum_{e_1,e_2\in V_m^{\text{2BORDER}}} \sum_{\substack{q_1\in\mathcal{P}_{e_1}^{\text{W}},\\ q_2\in\mathcal{P}_{e_2}^{\text{W}}}} \delta_{q_1}^{q_2} - \mu_4 \sum_{e,e'\in V_m^{\text{2BORDER}}} \delta_{e'}^{e} - \mu_5 \sum_{e,e'\in V_m^{\text{2BORDER}}} \sum_{\substack{q\in\mathcal{P}_e^{\text{W}},\\ q'\in\mathcal{P}_{e'}^{\text{B}}}} \delta_{q'}^{q}\Bigg). \quad (7)$$

The coefficients should be set carefully in order to define a meaningful objective. In general, μ_1 and μ_2 should be set large enough so that the two first terms, and thus bandwidth saving, are prioritized. The three last terms will help to select the solutions with the most topologically protect-able pairs of border nodes and the most disjoint intra-paths. Since working and backup intra-paths are relatively symmetrical in the Mapping, we can set $\mu_1 = \mu_2$ and $\mu_3 = \mu_5$. The coefficient μ_4 and μ_5 should be chosen so that the fifth term is always smaller than the increasing step of the fourth term in order to not act upon the maximization of the fourth term.

3.2. Exact solution for Mapping

The Mapping problem is complex since it looks for intra-paths for multiple pairs of border nodes and these intra-paths depend on each other. We use an Integer Linear Program (ILP) for modelling the optimal Mapping for each domain. Let us consider domain N_m.

Let h^e is the head border node and t^e is the tail border node of virtual link e then $e = (h^e, t^e)$. Let $x_{(u,v)}^{e,i}$ be the decision variable indicating if link (u, v) belongs to working intra-path $q_{e,i}^{\text{W}}$, indexed i, of border node pair $e \in V_m^{\text{2BORDER}}$:

$$x_{(u,v)}^{e,i} = \begin{cases} 1 & \text{if } (u,v) \in q_{e,i}^{\text{W}} \in \mathcal{P}_e^{\text{W}} \\ 0 & \text{otherwise} \end{cases} \qquad i = 1..n_e^{\text{W}}, \quad e \in V_m^{\text{2BORDER}}. \quad (8)$$

Let $y^{e,i}_{(u,v)}$ be the decision variable indicating if link (u,v) belongs to backup intra-path $q^{\mathrm{B}}_{e,i}$, indexed i, of border node pair $e \in V^{\mathrm{2BORDER}}_m$:

$$y^{e,i}_{(u,v)} = \begin{cases} 1 & \text{if } (u,v) \in q^{\mathrm{B}}_{e,i}, q^{\mathrm{B}}_{e,i} \in \mathcal{P}^{\mathrm{B}}_e \\ 0 & \text{otherwise} \end{cases}$$

$$i = 1..n^{\mathrm{B}}_e, \quad e \in V^{\mathrm{2BORDER}}_m. \tag{9}$$

3.2.1. Flow conservation constraint for working intra-paths

The flow conservation constraint for the working intra-path $q^{\mathrm{W}}_{e,i} \in \mathcal{P}^{\mathrm{W}}_e$ is:

$$\sum_u x^{e,i}_{(u,v)} - \sum_u x^{e,i}_{(v,u)} = \begin{cases} 1 & \text{if } v = h^e \\ 0 & \text{if } v \neq h^e, t^e \\ -1 & \text{if } v = t^e \end{cases},$$

$$v \in V_m, i = 1..n^{\mathrm{W}}_e, e \in V^{\mathrm{2BORDER}}_m. \tag{10}$$

In order to guarantee that $q^{\mathrm{W}}_{e,i}$ is a direct intra-path the following constraint is added:

$$x^{e,i}_{(u,v)} = 0 \text{ and } x^{e,i}_{(v,u)} = 0,$$

$$v \in V^{\mathrm{BORDER}}_m, u \in V_m, v \neq h^e, v \neq t^e,$$

$$i = 1..n^{\mathrm{W}}_e, e \in V^{\mathrm{2BORDER}}_m. \tag{11}$$

Although dummy loop does not affect neither the feasibility nor the value of the optimal solution, they are not desirable because they lead to identical solutions in the practice. The following constraints eliminate the loops.

$$\sum_u x^{e,i}_{(u,v)} \begin{cases} \leq 1, & \text{if } v \in V_m, v \neq h^e \\ = 0, & \text{if } v = h^e \end{cases} \quad e \in V^{\mathrm{2BORDER}}_m. \tag{12}$$

3.2.2. Flow conservation constraint for backup intra-paths

Similar to the flow conservation for working intra-paths, the following constraints apply for each backup intra-path:

$$\sum_u y^{e,i}_{(u,v)} - \sum_u y^{e,i}_{(v,u)} = \begin{cases} 1 & \text{if } v = h^e \\ 0 & \text{if } v \neq h^e, t^e \\ -1 & \text{if } v = t^e \end{cases},$$

$$v \in V_m, i = 1..n^{\mathrm{B}}_e, e \in V^{\mathrm{2BORDER}}_m. \tag{13}$$

$$y^{e,i}_{(u,v)} = 0 \text{ and } y^{e,i}_{(v,u)} = 0,$$
$$v \in V^{\text{BORDER}}_m, u \in V_m, v \neq h^e, v \neq t^e,$$
$$i = 1..n^{\text{B}}_e, e \in V^{\text{2BORDER}}_m. \quad (14)$$

$$\sum_u y^{e,i}_{(u,v)} \begin{cases} \leq 1, & \text{if } v \in V_m, v \neq h^e, \\ = 0, & \text{if } v = h^e. \end{cases}, e \in V^{\text{2BORDER}}_m. \quad (15)$$

3.2.3. Diversity condition

The following diversity condition forces the intra-paths in $\mathcal{P}^{\text{W}}_e$ to be distinct, so do the intra-paths in $\mathcal{P}^{\text{B}}_e$. Let us start with the diversity condition for working intra-paths. For each pair of working intra-paths $q^{\text{W}}_{e,i}, q^{\text{W}}_{e,j} \in \mathcal{P}^{\text{W}}_e$, $B^{q^{\text{W}}_{e,i} q^{\text{W}}_{e,j}}_{(u,v)}$ indicates whether one of them goes through link (u, v), i.e., $B^{q^{\text{W}}_{e,i} q^{\text{W}}_{e,j}}_{(u,v)}$ is equal to 1 if $q^{\text{W}}_{e,i}$ or $q^{\text{W}}_{e,j}$ goes through (u, v) and 0 otherwise, thus:

$$B^{q^{\text{W}}_{e,i} q^{\text{W}}_{e,j}}_{(u,v)} = \begin{cases} 0 & \text{if } (u,v) \notin q^{\text{W}}_{e,i}, q^{\text{W}}_{e,j} \\ 1 & \text{otherwise.} \end{cases},$$
$$i \neq j, e \in V^{\text{2BORDER}}_m. \quad (16)$$

which is linearized as follows:

$$\frac{1}{2}\left(x^{e,i}_{(u,v)} + x^{e,j}_{(u,v)}\right) \leq B^{q^{\text{W}}_{e,i} q^{\text{W}}_{e,j}}_{(u,v)} \leq x^{e,i}_{(u,v)} + x^{e,j}_{(u,v)} \quad (17)$$
$$B^{q^{\text{W}}_{e,i} q^{\text{W}}_{e,j}}_{(u,v)} \in \{0, 1\}.$$

Two intra-paths q_1, q_2 of the same pair of border nodes must have at least one merging and one switching points such as v in cases (a), (b), (c) (d) or (e) of Fig. 6. If case (b) occurs at a node, case (a) or (c) or (d) must occur at another node because q_1, q_2 have to join each other at the tail node of their virtual link. Similarly, if case (e) occurs at a node, case (a) or (c) or (d) must occur at some other nodes. Therefore, the diversity condition is satisfied if there exists a node v on q_1, q_2 so that at least one of cases (a), (c) or (d) occurs. That means, $\sum_u B^{q_1 q_2}_{(u,v)} = 2$. If v is neither a switching nor a merging point, i.e., cases (f), (g), (h) of Fig. 6, then $\sum_u B^{q_1 q_2}_{(u,v)} = 0$ or 1.

The diversity condition is thus expressed by:

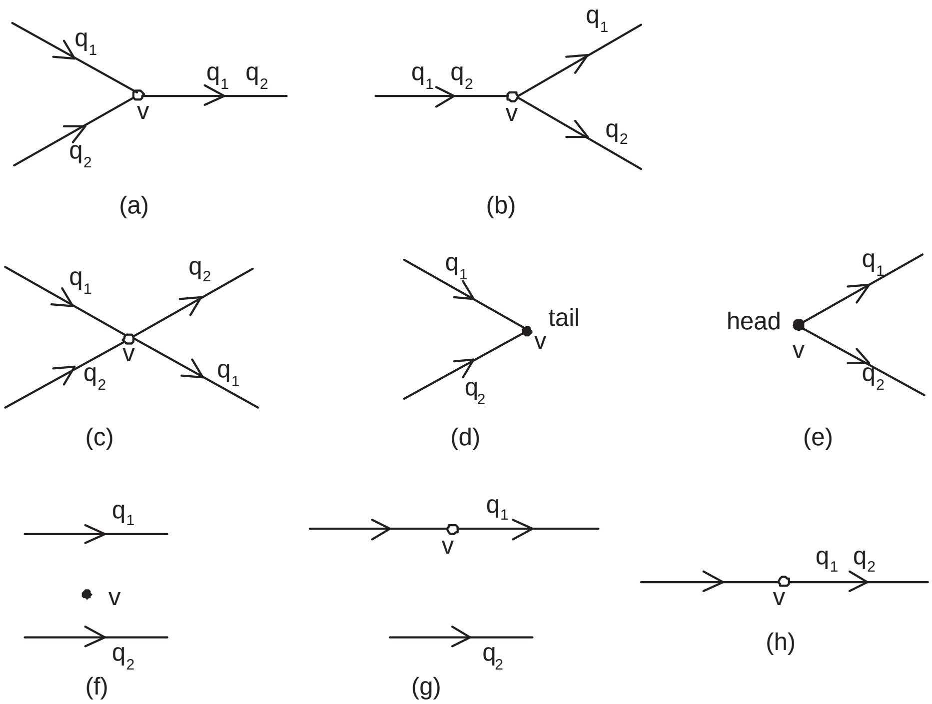

Figure 6: Possible cases for node v with respect to two intra-paths q_1 and q_2 of the same pair of border nodes. Cases (a), (b), (c), (d), (e): node v is a merging or switching point. Cases (f), (g), (h): node v is not a switching or merging point.

$$\exists v \in V_m, \sum_u B^{q_1 q_2}_{(u,v)} = 2. \tag{18}$$

For linearising (18), we introduce $r_v^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}} \in \{0,1\}$ as the decision variable which takes the value 1 if $\sum_u B^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}}_{(u,v)} = 2$ and 0 otherwise. Hence, $r_v^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}} = 1$ if $q^{\mathrm{W}}_{e,i}$ cuts $q^{\mathrm{W}}_{e,j}$ at v. Then, the diversity condition (18) is equivalent of the following two constraints:

$$\frac{1}{2} \sum_u B^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}}_{(u,v)} \geq r_v^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}} \geq \sum_u B^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}}_{(u,v)} - 1,$$
$$v \in V_m, i, j = 1..n^{\mathrm{W}}_e, i \neq j, e \in V_m^{\mathrm{2BORDER}} \tag{19}$$

$$\sum_{v \in V_m} r_v^{q^{\mathrm{W}}_{e,i} q^{\mathrm{W}}_{e,j}} \geq 1, \quad i, j = 1..n^{\mathrm{W}}_e, i \neq j, e \in V_m^{\mathrm{2BORDER}}. \tag{20}$$

The diversity among backup intra-paths can be defined similarly.

3.2.4. Disjointness between intra-paths

Let's first consider the disjointness between two direct working intra-paths $q^{\mathrm{W}}_{e_1,i} \in \mathcal{P}^{\mathrm{W}}_{e_1}, q^{\mathrm{W}}_{e_2,j} \in \mathcal{P}^{\mathrm{W}}_{e_2}$. Let:

$$A_v^{q^{\mathrm{W}}_{e_1,i} q^{\mathrm{W}}_{e_2,j}} = \sum_u x^{e_1,i}_{(u,v)} + x^{e_1,i}_{(v,u)} + x^{e_2,j}_{(u,v)} + x^{e_2,j}_{(v,u)},$$
$$v \in V_m, e_1, e_2 \in V_m^{\mathrm{2BORDER}}, i = 1..n^{\mathrm{W}}_{e_1}, j = 1..n^{\mathrm{W}}_{e_2}. \tag{21}$$

In Fig. 7, cases from (a) to (f) show the possible positions of a node v with respect to two intra-paths q_1, q_2 of two virtual links.

In cases (a), (b) (c): $A_v^{q_1 q_2} = 4$. In cases (d): $A_v^{q_1 q_2} = 2$. In case (e), v is on only one intra-path and is not a border node, then $A_v^{q_1 q_2} = 2$. In case (f), v does not belong to any intra-path, then $A_v^{q_1 q_2} = 0$. Therefore, the

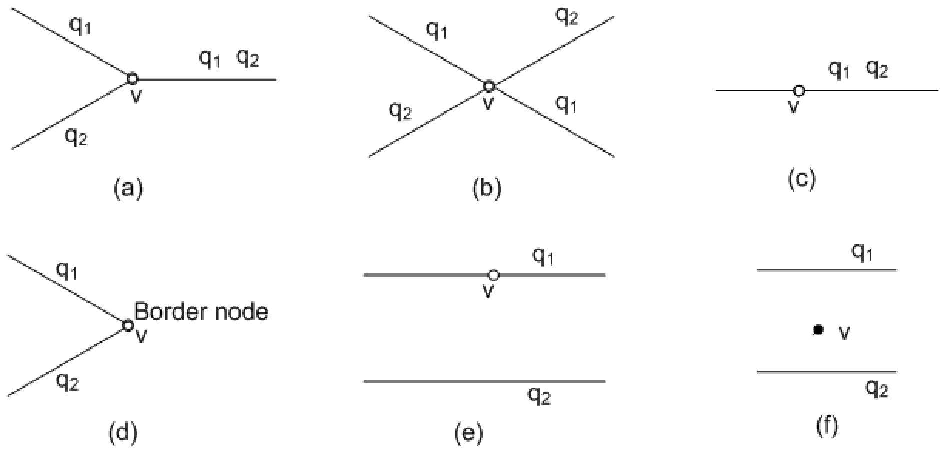

Figure 7: Positions of a node v with respect to two direct intra-paths of two pairs of border nodes regardless their directions.

disjointness between $q^{\mathrm{W}}_{e_1,i}$ and $q^{\mathrm{W}}_{e_2,j}$ is defined by:

$$\delta^{q^{\mathrm{W}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} = \begin{cases} 0 & \text{if } e, e_2 \text{ have a common borders node} \\ 0 & \text{if } \exists v \in V_m \setminus V^{\mathrm{BORDER}}_m : A_v^{q^{\mathrm{W}}_{e_1,i} q^{\mathrm{W}}_{e_2,j}} = 4 \\ 1 & \text{otherwise} \end{cases}$$

$$e_1, e_2 \in V^{2\mathrm{BORDER}}_m, i = 1..n^{\mathrm{W}}_{e_1}, j = 1..n^{\mathrm{W}}_{e_2}, \quad (22)$$

and it is equivalent to

$$\delta^{q^{\mathrm{W}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} \begin{cases} = & 0 \text{ if } e, e_2 \text{ have common end nodes, otherwise} \\ \leq & 2 - \frac{A_v^{q^{\mathrm{W}}_{e_1,i} q^{\mathrm{W}}_{e_2,j}}}{2}, \forall v \in V_m \setminus V^{\mathrm{BORDER}}_m \end{cases} \quad (23)$$

with

$$\delta^{q^{\mathrm{W}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} \in \{0, 1\}.$$

The disjointness between a direct working and a direct backup intra-path is defined similarly. Let the two intra-paths be $q^{\mathrm{W}}_{e_1,i} \in \mathcal{P}^{\mathrm{W}}_{e_1}$ and $q^{\mathrm{B}}_{e_2,j} \in \mathcal{P}^{\mathrm{B}}_{e_2}$. We need to substitute $A_v^{q^{\mathrm{W}}_{e_1,i} q^{\mathrm{W}}_{e_2,j}}$ by $A_v^{q^{\mathrm{W}}_{e_1,i} q^{\mathrm{B}}_{e_2,j}}$ in (23) where:

$$A_v^{q^{\mathrm{W}}_{e_1,i} q^{\mathrm{B}}_{e_2,j}} = \sum_u x^{e_1,i}_{(u,v)} + x^{e_1,i}_{(v,u)} + y^{e_2,j}_{(u,v)} + y^{e_2,j}_{(v,u)}. \quad (24)$$

3.2.5. Topological protect-ability between two pairs of border nodes

According to Definition 2, a pair of border nodes e_1 is topologically protect-able by another pair of border nodes e_2 if there exists a working intra-path $q^{\mathrm{W}}_{e_1,i} \in \mathcal{P}^{\mathrm{W}}_{e_1}$ and a backup intra-path $q^{\mathrm{B}}_{e_2,i} \in \mathcal{P}^{\mathrm{B}}_{e_2}$ that are disjoint. Therefore, $\delta^{e_2}_{e_1}$ must satisfy:

$$\frac{1}{n^{\mathrm{W}}_{e_1} \times n^{\mathrm{B}}_{e_2}} \sum_{i=1}^{n^{\mathrm{W}}_{e_1}} \sum_{j=1}^{n^{\mathrm{B}}_{e_2}} \delta^{q^{\mathrm{B}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} \leq \delta^{e_2}_{e_1} \leq \sum_{i=1}^{n^{\mathrm{W}}_{e_1}} \sum_{j=1}^{n^{\mathrm{B}}_{e_2}} \delta^{q^{\mathrm{B}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} \tag{25}$$

$$\delta^{e_2}_{e_1} \in \{0, 1\}.$$

3.2.6. Objective function

The objective function becomes:

$$\begin{aligned} \min\Bigg(\quad & \mu_1 \sum_{e \in V^{2\mathrm{BORDER}}_m} \sum_{i=1}^{n^{\mathrm{W}}_e} \sum_{(u,v) \in L_m} \frac{x^{e,i}_{(u,v)}}{c^{\mathrm{res}}_{(u,v)}} \\ + \quad & \mu_2 \sum_{e \in V^{2\mathrm{BORDER}}_m} \sum_{i=1}^{n^{\mathrm{B}}_e} \sum_{(u,v) \in L_m} \frac{y^{e,i}_{(u,v)}}{c^{\mathrm{res}}_{(u,v)}} \\ - \quad & \mu_3 \sum_{e_1, e_2 \in V^{2\mathrm{BORDER}}_m} \sum_{i=1}^{n^{\mathrm{W}}_{e_1}} \sum_{j=1}^{n^{\mathrm{W}}_{e_2}} \delta^{q^{\mathrm{W}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} \\ - \quad & \mu_4 \sum_{e_1, e_2 \in V^{2\mathrm{BORDER}}_m} \delta^{e_2}_{e_1} \\ - \quad & \mu_5 \sum_{e_1, e_2 \in V^{2\mathrm{BORDER}}_m} \sum_{i=1}^{n^{\mathrm{W}}_{e_1}} \sum_{j=1}^{n^{\mathrm{B}}_{e_2}} \delta^{q^{\mathrm{B}}_{e_2,j}}_{q^{\mathrm{W}}_{e_1,i}} \Bigg). \end{aligned} \tag{26}$$

The coefficients $\mu_1, \mu_2, \mu_3, \mu_4$ and μ_5 should be carefully chosen as already discussed in Section 3.1.

3.3. Heuristic solution for Mapping

The ILP model is complex and requires a very high computational effort for solving it. This section presents a more time efficient greedy heuristic. The main idea of the heuristic is as follows. We do not consider all possible intra-paths but only a subset $\mathcal{P}^{\mathrm{CAN}}_e \subset \mathcal{P}_e$ of n^{CAN}_e intra-path candidates for each pair of border nodes e. Of course, $n^{\mathrm{CAN}}_e \geq n^{\mathrm{W}}_e, n^{\mathrm{CAN}}_e \geq n^{\mathrm{B}}_e$. For satisfying Criteria 1 and 2, $\mathcal{P}^{\mathrm{CAN}}_e$ is the set of shortest intra-paths weighted by link residual capacities.

Again, the Mapping is performed independently in each domain. It is started with an empty list of intra-paths for each pair of border nodes. The list of pair of border nodes of the domain in consideration is browsed. For each pair of border nodes, we try first to find several working intra-paths so that they increase the most the number of pairs of border nodes that can topologically protect this pair of border nodes. Next, amongst the found working intra-paths, we select the one that is disjoint with the largest number of existed working intra-paths. Then a backup intra-path is also identified for the pair of border nodes as the intra-path that increases the most the number of topologically protect-able pairs of border nodes. The next pair of border nodes will be considered in the same way. When all pairs of border nodes are visited, another round is started again and again until each pair of border nodes receives the required number of potential working and backup intra-paths. Pseudo-code in Alg. 1: Greedy_mapping(N_m) details the algorithm.

3.4. Mapping refresh

After a certain time of operation, links on potential intra-paths may be running out of residual bandwidth. A new Mapping should be run in order to find new potential intra-paths with more residual bandwidth. Let threshold ϵ^{res} be the smallest amount of bandwidth remaining in each intra-path before the Mapping process should be run again. In order to avoid blocking due to link saturation, ϵ^{res} must not be smaller than the smallest requested bandwidth and is not necessary to be greater than the largest requested bandwidth.

Therefore, as soon as the residual capacity of an potential intra-path gets to ϵ^{res}, the Mapping should be refreshed for the domain containing the intra-path.

4. Routing

4.1. Mathematical formulation

The objective of the Routing phase is, to find working and backup segments for a new incoming request so that their total bandwidth cost is minimized. Let d be the requested bandwidth. As being stated in Section 2, the Routing problem is solved easily by using a single domain routing solution on mapped network. Single domain routing solutions for shared path protection can be found in [8], [26], etc. Those for shared segment protection can be found in [4], [13], etc.. The inputs of these algorithms are network topology,

Algorithm 1 Greedy_mapping(N_m)

```
for all e ∈ V_m^{2BORDER} do
  P_e^{CAN} = set of n^{CAN} shortest intra-paths weighted by residual capacity.
end for
while ∃e ∈ V_m^{2BORDER} so that |P_e^W| < n_e^W and |P_e^B| < n_e^B do
  {—Some pairs of border nodes have not received enough potential intra-paths —}
  for all e ∈ V_m^{2BORDER} do
    if |P_e^W| < n_e^W then
      {—Select an intra-path for e if its set of potential working intra-paths is not full—}
      for all q ∈ P_e^{CAN} do
        dj_q ← Number of pairs of border nodes that can newly topologically protect e thanks to q
      end for
      S_e ← Set of n intra-paths that have the highest dj_q
      q ← The intra-path in S_e that is disjoint with the largest number of working intra-paths in ⋃_{e_1 ∈ V_m^{2BORDER}} P_{e_1}^W
      P_e^W = P_e^W ∪ {q}
    end if
    if |P_e^B| < n_e^B then
      {—Select an intra-path for e if P_e^B is not full—}
      for all q ∈ P_e^{CAN} do
        dj_q ← Number of pairs of border nodes that are newly topologically protect-able by e thanks to q
      end for
      q ← The intra-path that has the highest dj_q
      P_e^B = P_e^B ∪ {q}
    end if
  end for
end while
```

working cost of each network link and backup cost of each network link for protecting another link. In this section, we present how to compute those costs for links of the mapped network. It worth to note that, we take care of node protection.

Let q (resp. q') be a working (resp. backup) intra-path/virtual edge that is considered for the new incoming request. Let us assume that a bandwidth unit on a physical link has a unit cost and the bandwidth cost of a segment is the sum of the bandwidth costs of its links.

c_ℓ^{res} residual capacity of physical link ℓ.

$B_{q'}$ bandwidth reserved by backup segments going through the entire virtual edge q'. Be aware that $B_{q'}$ may differ from the total backup bandwidth reserved on a physical link of q'. (see an example in Fig. 4).

$B_{q'}^v$ sum of the bandwidth requested by the connections of which a backup segment goes through q' and the corresponding working segment goes through node v. Those backup segments cannot share bandwidth amongst them because they will be activated simultaneously when v fails. Their backup bandwidth is not profitable for a backup segment of the new incoming request if this segment goes through q' and protects a working segment going through v.

$B_{q'}^q$ sum of bandwidth requested by the connections of which a backup segment goes through q' and the corresponding working segment goes through q.

$\delta_{q'}^q$ disjointness between two intra-paths q, q'. It is set to 1 if q and q' are node disjoint and 0 otherwise.

α_q total cost of the working bandwidth used by the new incoming request along virtual edge q.

$\beta_{q'}^q$ backup bandwidth cost of the new incoming request along virtual edge q' for protecting virtual edge q against any single link or node failure. Note that $\beta_{q'}^q$ is the additional bandwidth that the new incoming request needs on q' excluding the fraction of shareable backup bandwidth it can benefit from $B_{q'}$.

Definition 3. *The residual capacity of a virtual edge is the maximum bandwidth that we can route along it.*

Theorem 1. *The residual capacity of virtual edge q is the minimum of the residual capacity on each of its physical links:*

$$\gamma_q = \min_{\ell \in q} c_\ell^{\text{res}}. \tag{27}$$

Proof. Indeed, the smallest residual capacity on q is $\min_{\ell \in q} c_\ell^{\text{res}}$. We can always route an amount of bandwidth $\min_{\ell \in q} c_\ell^{\text{res}}$ over q because every link along q has at least this amount of available capacity. On the other hand, there is at least one link of q whose residual capacity is $\min_{\ell \in q} c_\ell^{\text{res}}$, thus we cannot route more than this bandwidth over q. As a result, $\gamma_q = \min_{\ell \in q} c_\ell^{\text{res}}$

We are now going to identify the working and backup costs of virtual edges. Unlike RaM, all costs will be computed exactly in MaR. Since working segments do not share any bandwidth, each working segment uses exactly bandwidth d over each of its physical links. The working cost of the new incoming request on virtual edge q amounts to:

$$\alpha_q = \begin{cases} \|q\| \times d & \text{if } d \leq \gamma_q \\ \infty & \text{otherwise.} \end{cases} \tag{28}$$

Theorem 2. *If the bandwidth reserved on a backup segment is sufficient for protecting its working segment against a failure on a node, it is also sufficient for protecting the same working segment against a failure on a link adjacent to the node.*

Proof. Let us denote the backup segment by p' and the working segment by p. Let v be a node of p and e be an adjacent link of v. Let $\mathbb{P}_{p'}^v$ and $\mathbb{P}_{p'}^e$ be respectively the sets of working segments going through v and e such that their backup segments go through p'. Hence, $\mathbb{P}_{p'}^e \subseteq \mathbb{P}_{p'}^v$. Assume that the backup bandwidth reserved on p' is sufficient against a failure at node v, it means that this bandwidth is sufficient for activating simultaneously all backup segments in $\mathbb{P}_{p'}^v$. The backup bandwidth is thus sufficient for activating simultaneously all backup segments in $\mathbb{P}_{p'}^e$. In other words, it is sufficient for recovering the failure at e.

This theorem shows that in order to protect a working segment against failures on nodes and links, we only need to protect nodes and then links will be automatically protected.

Theorem 3. *The backup cost of the new incoming request on virtual edge q' for protecting a virtual edge q against any single link or node failure is:*

$$\beta^q_{q'} = \begin{cases} \|q'\| \times (\max_{v\in q} B^v_{q'} + d - B_{q'}) & \text{if } \delta^q_{q'} = 1 \text{ and} \\ & 0 \le \max_{v\in q} B^v_{q'} + d - B_{q'} \le \gamma_{q'} \\ 0 & \text{if } B_{q'} - \max_{v\in q} B^v_{q'} \ge d \\ \infty & \text{otherwise.} \end{cases} \tag{29}$$

Proof. when q, q' are not disjoint, i.e., $\delta^q_{q'} = 0$, they both fail upon a single failure at a common link or node, therefore $\beta^q_{q'} = \infty$. Otherwise, let us consider the backup bandwidth needed by the new incoming request on a physical link of q' in order to protect q against a failure on node $v \in q$. Within the existing backup bandwidth $B_{q'}$ on q', $B^v_{q'}$ is non shareable for covering a failure at v. The remaining bandwidth is shareable and amounts to $B_{q'} - B^v_{q'}$ for every physical link of q'. Thus, the additional bandwidth that the new incoming request needs on each physical link of q' for protecting v is: $B^v_{q'} + d - B_{q'}$. In the single failure context, only one node can fail at a time. Hence, the additional backup bandwidth needed on a physical link of q' for protecting q against any single failure is: $\max_{v\in q}(B^v_{q'} + d - B_{q'})$. As the cost of q' is proportional to its length, we deduce the formula (29).

At this point, we have successfully identified the working and backup costs of a virtual edge in the mapped network. Single domain Shared Path Protection or Shared Segment Protection can be used in this network with the above defined working and backup costs for finding the final working and backup paths/segments.

4.2. An exact and scalable solution for computing the backup cost of a virtual edge

The computation of backup cost $\beta^q_{q'}$ as expressed in (29) requires the knowledge of $B^v_{q'}$ for each node v and intra-path q'. It is an intra-domain information which changes dynamically after each routing. Therefore, maintaining all $B^v_{q'}$ up-to-date is a non-scalable requirement.

We propose a more scalable method for computing $\beta^q_{q'}$ based on maximal Shared Risk Virtual Link Groups concept. The idea is that in each domain, there exist some critical nodes which belong to many intra-paths. The protection of these nodes can be sufficient for protecting some other nodes (see Th. 4 below). Therefore, the backup cost for protecting an intra-path can be deduced from the backup cost for protecting some given critical nodes.

Shared risk link group has been proposed as a fundamental concept in failure management. A Shared risk link group is a group of network links that share a common physical resource (cable, conduit, node or substructure) whose failure will cause the failure of all links in the group [27]. The working path and its backup path must not use links in the same Shared risk link group. In this paper, the notion Shared Risk Virtual Edge Group is defined in similar principal with Shared risk link group but for virtual edges.

Definition 4. *A Shared Risk Virtual Edge Group (*SRVEG*) at a node v is a set of virtual edges that all fail when the node v fails.*

The $\mathrm{SRVEG}(v)$ is actually the set of virtual edges (intra-paths) going through v. SRVEG has the following characteristic:

Theorem 4. *Let* $\mathrm{SRVEG}(v_i)$ *and* $\mathrm{SRVEG}(v_j)$ *be two* SRVEG*s so that* $\mathrm{SRVEG}(v_i) \subseteq \mathrm{SRVEG}(v_j)$*, then*

$$B_{q'}^{v_i} \leq B_{q'}^{v_j}. \tag{30}$$

In other words, if all the intra-paths that go through node v_i, go also through another node v_j, the backup bandwidth reserved on the intra-path for protecting v_i does not exceed the backup bandwidth reserved on the same intra-path for protecting v_j, see Fig. 8 for an illustration.

Proof. Since $\mathrm{SRVEG}(v_i) \subseteq \mathrm{SRVEG}(v_j)$ then:

$$\mathrm{SRVEG}(v_j) = \mathrm{SRVEG}(v_i) \cup \Big(\mathrm{SRVEG}(v_j) \setminus \mathrm{SRVEG}(v_i) \Big).$$

Thus:

$$\sum_{q \in \mathrm{SRVEG}(v_j)} B_{q'}^{q} = \sum_{q \in \mathrm{SRVEG}(v_i)} B_{q'}^{q} + \sum_{q \in \mathrm{SRVEG}(v_j) \setminus \mathrm{SRVEG}(v_i)} B_{q'}^{q}.$$

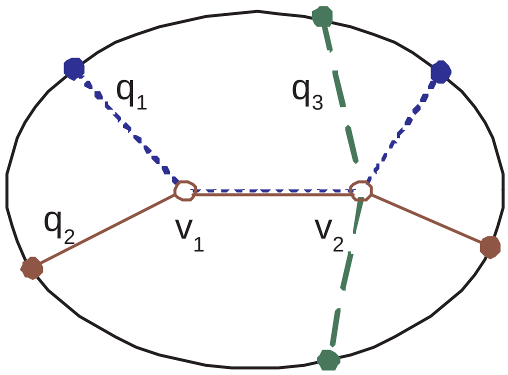

Figure 8: $\mathrm{SRVEG}(v_1) = \{q_1, q_2\} \subset \mathrm{SRVEG}(v_2) = \{q_1, q_2, q_3\}$ because all intra-paths going through v_1 go also through v_2.

Consequently,

$$\sum_{q \in \mathrm{SRVEG}(v_j)} B^q_{q'} \geq \sum_{q \in \mathrm{SRVEG}(v_i)} B^q_{q'}.$$

From the definition of $B^v_{q'}$, we have $B^v_{q'} = \sum_{q \in \mathrm{SRVEG}(v)} B^q_{q'}$. Thus,

$$B^{v_j}_{q'} \geq B^{v_i}_{q'}.$$

Definition 5. *A* SRVEG *is maximal if it is not contained in another* SRVEG*.*

From Th. 4 we deduce the following theorem.

Theorem 5. *Let q be a sub-path and v_j be a node on q such that* $\mathrm{SRVEG}(v_j)$ *is maximal and denoted by* $\mathrm{SRVEG}^{\mathrm{MAX}}(v_j)$*. We have:*

$$\max_{v \in q} B^v_{q'} = \max_{v_j : q \in \mathrm{SRVEG}^{\mathrm{MAX}}(v_j)} B^{v_j}_{q'}. \tag{31}$$

Proof. First note that $v_j : q \in \mathrm{SRVEG}^{\mathrm{MAX}}(v_j)$ is the formal expression of the fact that v_j is in q and $\mathrm{SRVEG}(v_j)$ is maximal. The proof is now as follows.

On the one hand, since v_j is in q, the set $\{v_j : q \in \mathrm{SRVEG}^{\mathrm{MAX}}(v_j)\}$ is a subset of the set $\{v : v \in q\}$. Therefore,

$$\max_{v \in q} B^v_{q'} \geq \max_{v_j : q \in \mathrm{SRVEG}^{\mathrm{MAX}}(v_j)} B^{v_j}_{q'}. \tag{32}$$

On the other hand, for all $v \in q$, there exists a maximal SRVEG that contains or is equal to SRVEG(v). Let SRVEG$^{\text{MAX}}(v_k)$ be such a SRVEG, then SRVEG$(v) \subseteq$ SRVEG$^{\text{MAX}}(v_k)$. According to Th. 4, $B^v_{q'} \leq B^{v_k}_{q'}$. In addition, since $q \in$ SRVEG(v) then $q \in$ SRVEG$^{\text{MAX}}(v_k)$ and thus v_k is in q. Consequently, for each $v \in q$ there exists $v_k : q \in$ SRVEG$^{\text{MAX}}(v_k)$ such that and $B^v_{q'} \leq B^{v_k}_{q'}$. Hence,

$$\max_{v \in q} B^v_{q'} \leq \max_{v_k : q \in \text{SRVEG}^{\text{MAX}}(v_k)} B^{v_k}_{q'}. \tag{33}$$

Inequalities (32) and (33) lead to (31).

Th.5 provides a way to compute $\max_{v \in q} B^v_{q'}$. Let v_1, v_2 be two border nodes of virtual edge q and N_m be the domain containing q. Then:

$$\max_{v \in q} B^v_{q'} = \max \left\{ B^{v_1}_{q'}, B^{v_2}_{q'}, \max_{\substack{v_j \in V_m \setminus V_m^{\text{BORDER}}: \\ q \in \text{SRVEG}^{\text{MAX}}(v_j)}} B^{v_j}_{q'} \right\}. \tag{34}$$

The third term of the maximum in (34) relates only to non-border nodes $v_j \in V_m \setminus V_m^{\text{BORDER}}$ while the first and second terms take care of border nodes v_1, v_2. We do not need to consider the other border nodes since q is a direct intra-path. Note that $B^{v_1}_{q'}$ and $B^{v_2}_{q'}$ can be easily identified by looking at the virtual edges going through v_1 and v_2.

Equation (34) holds also when q is an inter-domain link as we then have $V_m \setminus V_m^{\text{BORDER}} = \emptyset$.

In substituting $\max_{v \in q} B^v_{q'}$ in (29) by the right hand-side of (34), we obtain a scalable formula for computing the backup cost $\beta^q_{q'}$.

In conclusion, in order to identify the backup cost of an intra-path for protecting another intra-path, we only need to compute the backup cost for protecting the border nodes of the intra-path and those for protecting some non-border nodes whose SRVEG is maximal. Such non-border node can be identified easily based on domain topology. They are usually the clue points in the domain.

5. Scalability discussion

Since the Mapping step is performed independently within each domain, it does not encounter any scalability difficulty. Let us discuss the scalability issue in the Routing process.

When a new request comes in, the source border node is responsible for identifying working and backup segments for the request by performing the Routing process. First of all, it needs to compute the working and backup costs associated with each virtual edge. According to (28), the required parameters for computing the working cost of each virtual edge q are: $\|q\|, \gamma_q$. According to (29) and (34), the required parameters for computing the backup cost of virtual edge q (q' in the formula) in order to protect all the other virtual edges q^{W} (q in the formula) are: $\|q\|, B_q^v$ for all $v \in V^{\text{BORDER}}$, B_q, γ_q, B_q^v for all non-border nodes v whose SRVEG is maximal.

Let E^{VEDGE} be the set of virtual edges in the mapped network. Each border node should store the parameters categorized as follows:

Cat.A : $\|q\|, B_q$ for all $q \in E^{\text{VEDGE}}$;

Cat.B : B_q^v for all $v \in V^{\text{BORDER}}$ and for all $q \in E^{\text{VEDGE}}$;

Cat.C : All non-border SRVEG$^{\text{MAX}}$ in every domain as well as their associated internal nodes v;

Cat.D : B_q^v for all internal nodes v that are associated with the non-border SRVEG$^{\text{MAX}}$ of Cat.C, and for all $q \in E^{\text{VEDGE}}$;

Cat.E : γ_q for all $q \in E^{\text{VEDGE}}$.

These parameters should be kept up-to-date and exchanged between border nodes, for example by using an extension of BGP in order to adapt with the newly introduced parameters. These exchanges provide the border nodes with the identical information for working and backup cost computing.

In Cat.A, values $\|q\|$, for all $q \in E^{\text{VEDGE}}$ do not need to be updated because they are constant unless the network topology changes. The parameter B_q for all $q \in E^{\text{VEDGE}}$ is easily managed in the mapped network by increasing or decreasing it by $\beta_q^{\pi_i}/\|q\|$ at each setting up or tearing down of a given connection request. In Cat.B, B_q^v for all $v \in V^{\text{BORDER}}$, $q \in E^{\text{VEDGE}}$ can also be managed in a similar way except that the increasing and reduction are equal to d. In Cat.C, the non-border SRVEG$^{\text{MAX}}$s depend uniquely on the routes of virtual edges so they are not changed until the next Mapping. Consequently, there is no need to exchange often these parameters. The experimental results in Section 6.2 will show that the number of non-border SRVEG$^{\text{MAX}}$ is quite small, therefore the number of B_q^v in Cat.D is also small.

Abbreviation	Description
MaR-O	MaR approach with the optimal mapping.
MaR-G	MaR approach with the greedy heuristic mapping.
GROS	GROS of RaM approach proposed in [25] for OSSP.
DYPOS	DYPOS of RaM approach proposed in [25] for OSSP.
Opt	Optimal single domain OSSP solution proposed in [4].
NoShare	Overlapping segment protection without bandwidth sharing.

Table 1: Abbreviation of different approaches

The values of B_q^v in Cat.B and Cat.D for a given virtual edge q can be stored at a border node of q. The values of B_q^v are required only for the computations of backup segments whose working path is previously identified in most routing algorithms used for the Routing step. In a backup segment computation, we can collect B_q^v for all v in the working path by sending a signalling message, along the working path and get those B_q^v back with returning message.

It now remains the values γ_q of Cat.E which need to be kept track for each virtual edge of E^{VEDGE}. Indeed, while the residual capacity on every physical link that participates in q is not smaller than the maximal requested bandwidth, we know that the residual capacity γ_q of virtual edge q is sufficient for any new request and does not need to be updated. Otherwise, γ_q needs to be recalculated exactly by using (27). The calculation impairs the most the scalability of our solution. however it is performed uniquely when the network is nearly saturated.

In summary, most of the information required in the routing of MaR is per virtual edge and is managed at the mapped level. The quantity of required internal domain information is small. The scalability is thus preserved.

6. Experimental results

The experiments are performed for OSSP scheme. In order to evaluate MaR approach, we compare its results with the results of several other approaches. The results are denoted according to approach names as in Table 1.

Since the role of working and backup intra-paths are symmetric in the Mapping, we assume that both working and backup traffic use a common set $\mathcal{P}_e^{\text{WB}}$ of n_e^{WB} potential intra-paths. Therefore, the model for MaR-O was implemented by removing from the original one the constraints and the objective terms related to backup intra-paths, i.e., the terms weighted by μ_2 and μ_5. As discussed in Section 3.1, bandwidth saving needs to be prioritized therefore, μ_1 is set large and μ_3 is set small enough so that the smallest increasing in term associated with μ_1 is still larger than any value of the term associated with μ_3. Consequently, in these experiments, the coefficients are set as follows: $\mu_1 = \max_{\ell \in L_m} c_\ell^{\text{res}}$, $\mu_3 = \frac{1}{(n^{\text{W}} \times |V_m^{2\text{BORDER}}|)^2}$ and $\mu_4 = 1$. The implemented model for MaR-G is deduced from the original one by removing the computations for backup intra-paths.

For both MaR-O and MaR-G, the number of needed intra-paths per pair of border nodes is $n^{\text{W}} = 2$ and the number of intra-path candidates for MaR-G is $n^{\text{CAN}} = 4$. The Mapping step of MaR-O is solved using Cplex [28]. Opnet Modeler [29] is used to implement the Routing step of both MaR-O and MaR-G. Note that in MaR-O, MaR-G, GROS and DYPOS, working and backup segment lengths are limited by thresholds l^{W} and l^{B} respectively.

6.1. Mapping evaluation

The greedy mapping of MaR-G is compared with the optimal mapping of MaR-O on 2 large multi-domain networks: LARGE-5 and LARGE-8. LARGE-5 is built using 5 real optical networks: EON [30] (29 nodes, 39 edges), RedIris [31] (19, 31), Garr [32] (15, 24), Renater [33] (18, 23), SURFnet [34] (25, 34). LARGE-8 is generated using the Transit-Stub model of GT-ITM [35], a well known multi-domain network generator. The network contains 8 domains, each one has on average 4 neighboring domains in order to reflect faithfully the Internet interconnections [36]. The numbers of nodes and links of each domain are: $(20, 53)$, $(20, 29)$, $(21, 48)$, $(22, 41)$, $(18, 36)$, $(20, 44)$, $(17, 27)$, $(22, 47)$.

For comparing MaR-G and MaR-O, we compute the objective values (7) of the mapping as well as each of its terms with those obtained from MaR-G and MaR-O. Table 2 gives the relative gaps between the values obtained from MaR-G and those obtained from MaR-O. The gap for the overall objective is in the last column of the table. Currently, the comparison is only made on LARGE-5 due to the too high computational effort for solving MaR-O in LARGE-8.

Most differences between MaR-G and MaR-O are found in the intra-path costs. Since the intra-path cost is prioritized in the objective function,

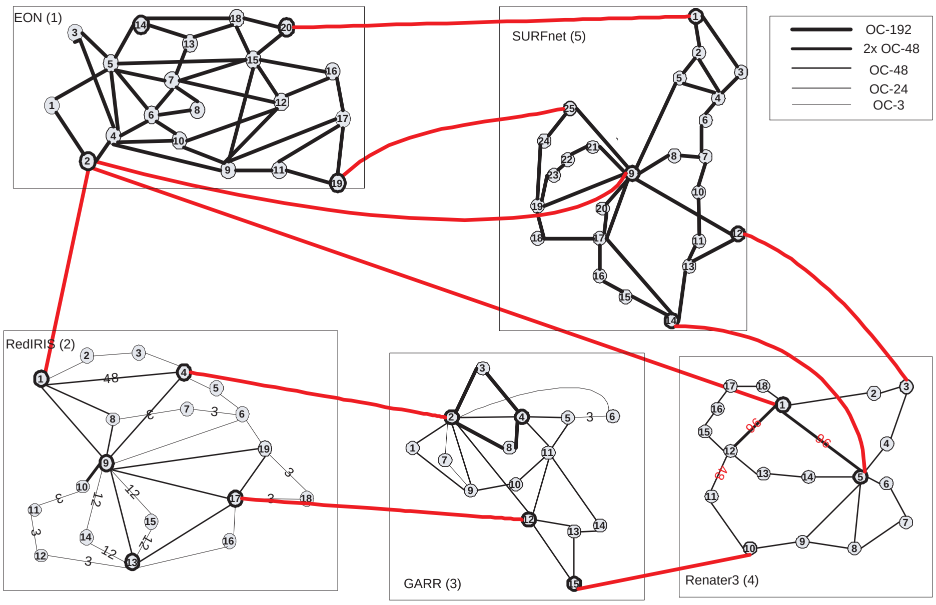

Figure 9: Multi-domain network LARGE-5

Domains	(μ_1) cost	($-\mu_3$) dj_{ip}	($-\mu_4$) dj_{border}	obj
EON	-2,17	-33	0	2.52
RedIRIS	0	0	0	0
GARR	0	0	0	0
Renater	25	0	0	26.48
SURFnet	7,81	0	0	10.89

cost (%): relative gap on intra-path cost.
dj_{border} (%): relative gap on number of sets of two pairs of border nodes that are topologically protect-able one another.
dj_{ip} (%): relative gap on number of disjoint intra-paths.
obj (%): relative gap on the overall objective function.

Table 2: Relative gap of MaR-G vs. MaR-O in LARGE-5.

Domains	EON	RedIRIS	GARR	Renater	SURFnet	Total
Nb. org. SRVEG (MaR-O)	16	15	13	12	21	77
Nb. org. SRVEG (MaR-G)	12	16	13	16	22	79
Nb. SRVEG$^{\text{MAX}}$ (MaR-O)	8	6	5	4	6	29
Nb. SRVEG$^{\text{MAX}}$ (MaR-G)	7	6	5	4	6	28
Nb. adv. SRVEG (MaR-O)	4	1	1	0	1	7
Nb. adv. SRVEG (MaR-G)	3	1	1	0	1	6

Table 3: Number of SRGs in LARGE-5

the differences reflect clearly in the overall objective gap. However the intra-path cost gaps are generally small, leading to small overall objective gaps in most domains, except for Renater domain.

In summary, the mapping results of MaR-G are close to those of the optimal Mapping MaR-O, illustrating the efficiency of the proposed greedy Mapping.

6.2. Scalability in using non-border maximal SRVEG*s*

Section 4.2 and 5 show that only the non-border maximal SRVEGs need to be advertised amongst domains. The smaller is the number of advertised SRVEGs, the more scalable MaR is. Table 3 shows the number of SRVEGs that need to be advertised (denoted by adv.), the number of original SRVEGs which would be required in a domain (denoted by org.), and the number of maximal SRVEGs in LARGE-5. Most domains require either 1 or no SRVEG to be advertised.

In LARGE-8, domains are highly connected with more links and internal nodes in comparison to those of LARGE-5. This characteristic leads to a less drastic reduction of the number of SRVEGs (see Table 4). However, more than 68% SRVEGs are still eliminated. The number of SRVEGs to be advertised per domain remains small.

In conclusion, the use of only non-border maximal SRVEGs in backup cost computation leads to a significantly more scalable routing solution while maintaining the accuracy of the cost.

Domains	1	2	3	4	5	6	7	8	Total
Nb. org. SRVEGs	16	16	21	18	18	20	15	17	141
Nb. SRVEG$^{\text{MAX}}$	5	12	17	9	17	19	10	13	102
Nb. adv. SRVEGs	1	4	8	3	8	10	4	6	44

Table 4: Number of SRVEGs of LARGE-8 with M*a*R-G

6.3. Routing evaluation

In this section, the Routing step is evaluated together with the Mapping step through the final routing results. In R*a*M, the inter-domain routing is performed by greedy algorithm GROS or dynamic programming algorithm DYPOS, then another routing is performed in intra-domain networks. In M*a*R, the same dynamic programming solution as in DYPOS is used for the unique routing of both M*a*R-G and M*a*R-O. Both GROS and DYPOS restrict the lengths of the working and backup segments by two parameters l^{W} and l^{B} respectively.

Following evaluation metrics are used for the routing evaluation:

- The *working (resp. backup) network cost* is the total working (resp. backup) bandwidth used by all network links.
- The *Backup overhead* is the ratio between the total working and backup network cost and the smallest working network cost minus 1. This amounts to the backup bandwidth redundancy of a protection scheme. In other words, it represents the backup bandwidth saving level of a protection scheme. The smallest working network cost is the total cost of all working paths without protection when these paths are shortest paths.
- The *Overall blocking probability* is the percentage of the total rejected bandwidth out of the total requested bandwidth.

6.3.1. Comparison with optimal single domain OSSP solution

Due to the extremely high computational effort for Opt, the results of M*a*R-O, M*a*R-G, GROS and DYPOS are compared with those of Opt only on a small 5 domain network of 28 nodes with 70 submitted requests. All requests remain active in the network without tearing down. The Transit-Stub model of GT-ITM is used again for generating this network instance that we denote by SMALL-5 and illustrate in Fig. 10.

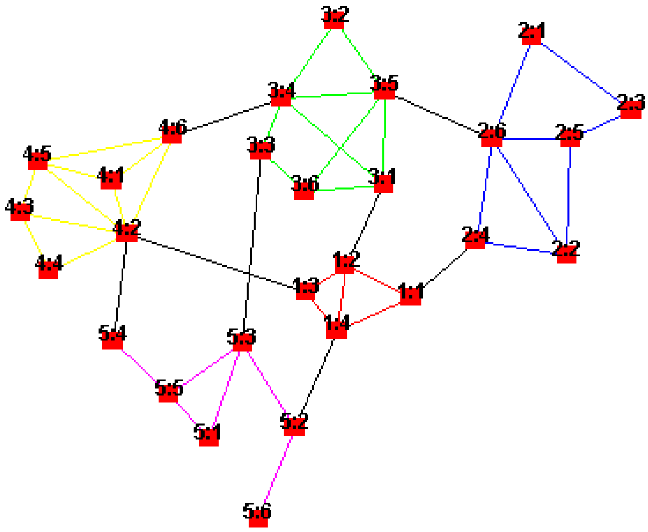

Figure 10: SMALL-5 network.

Fig. 11 depicts the backup overhead of different solutions in SMALL-5 when the working segment length threshold l^{W} varies from 2 to 5. Due to the small size of the network, there is no need for testing with larger l^{W}. For the same reason, the backup segment length constraint is removed. MaR-O and MaR-G outperform GROS and DYPOS for most values of l^{W}. When the constraint on working segment lengths is loose, i.e., l^{W} becomes large, MaR-O, MaR-G and Opt operate under similar conditions since no segment length constraint is required in Opt. In this case, MaR-O and MaR-G provide nearly identical backup overheads to Opt, revealing their high performances in bandwidth saving.

6.3.2. Backup overhead

An advantage of MaR over RaM is that it solves a single routing optimization problem instead of multiple optimizations in inter-domain and intra-domain routings. In addition, MaR does not suffer from the approximation of working and backup cost computations in inter-domain routing as in RaM. This allows MaR improving the quality of its solutions. The experiments in this section will show that MaR provides a smaller backup overhead and thus better bandwidth saving.

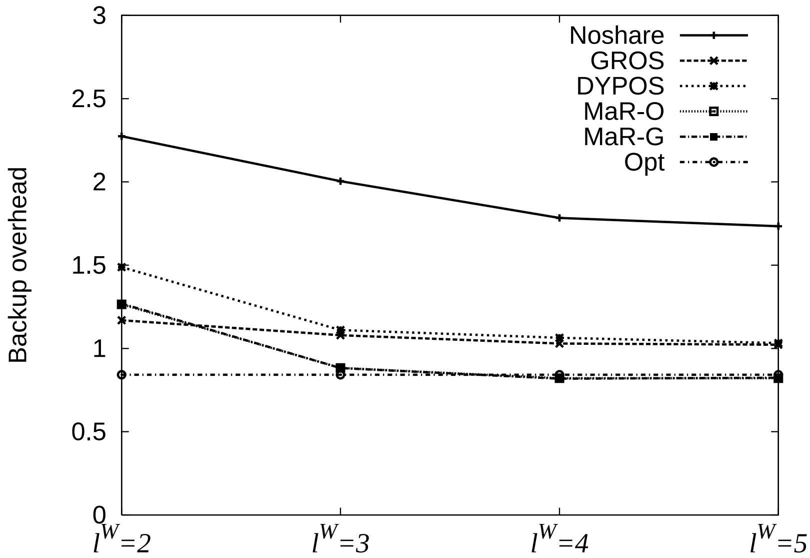

Figure 11: Comparison with Opt on Backup overhead in SMALL-5

We conducted experiments with an incremental traffic in large networks. The incremental traffic is generated by submitting subsequently 1000 connection requests to the network with all requests remaining active. Network links are uncapacitated in order to avoid the blocking cases which vary from one scheme to the other and thus would make the analysis more complex. Backup overhead is computed once after 1000 requests. No experiment is performed with MaR-O because of its high computational effort.

Fig. 12 and Fig. 13 depicts backup overheads of MaR-G, GROS, DYPOS and NoShare in LARGE-5 and LARGE-8 when working and backup segment length thresholds vary. Obviously, MaR-G, GROS and DYPOS give better backup overheads than NoShare. As expected, MaR-G provides generally a smaller backup overhead than GROS and DYPOS.

6.3.3. Blocking probability

The blocking probability is examined under dynamic traffic. Requests for connection arrive and tear down after a holding time. Requests arrive according to a Poisson process with rate $r = 1$ and with an exponentially distributed holding time with mean $h = 320$. There are, on average, 320 active connections in the network.

In general, MaR-G provides clearly smaller blocking probability than

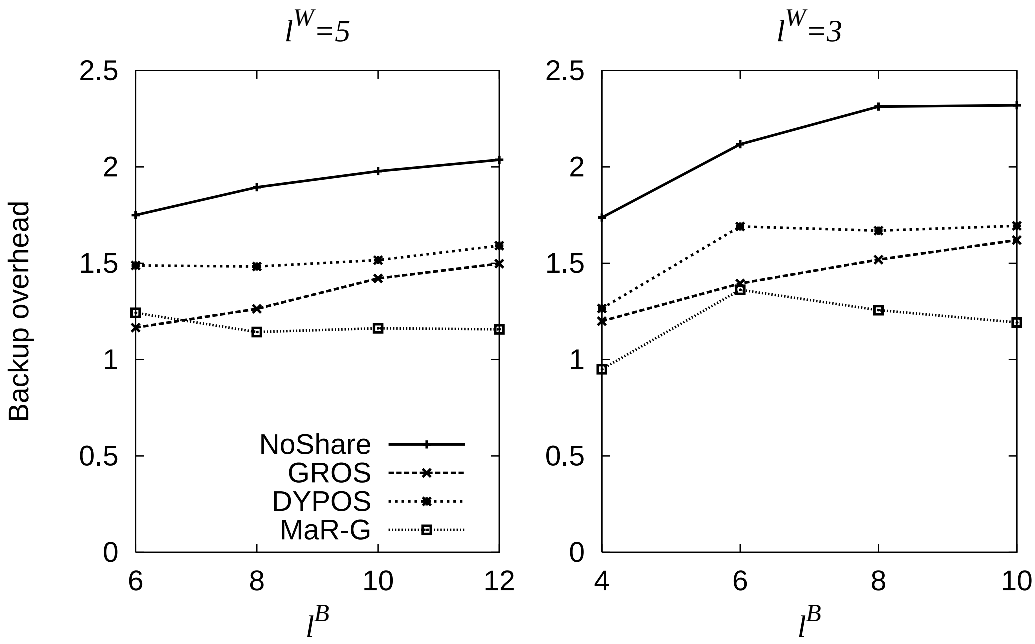

Figure 12: Backup overhead in LARGE-5

DYPOS, GROS and NoShare with $l^{\mathrm{W}} = 5$ in LARGE-5 (see Fig. 14) and in LARGE-8 (see Fig. 15). An insight in GROS and DYPOS reveals that most of their blocking is caused by bad guidances obtained from the inter-domain routing due to cost approximation and the impossibility of mapping virtual links in the intra-domain step to disjoint working and backup segments. MaR-G overcomes these weaknesses by using an unique routing based on precise working and backup costs of virtual edges as well as their disjointness indexes δ.

However, we observe from the results on both backup overhead and blocking probability that when segment lengths are highly limited, i.e., $l^{\mathrm{W}} = 3$ or small l^{B}, MaR-G sometimes looses its advantage. On the one hand, it is more difficult for MaR-G to build a solution satisfying segment length constraints from the restricted number of potential intra-paths ($n^{\mathrm{W}} = 2$) than GROS and DYPOS which have no restriction on the selection of intra-paths. On the other hand, DYPOS or GROS applies segment length constraints in the inter-domain routing which is based on approximation and thus the final results may not satisfy the segment length constraints.

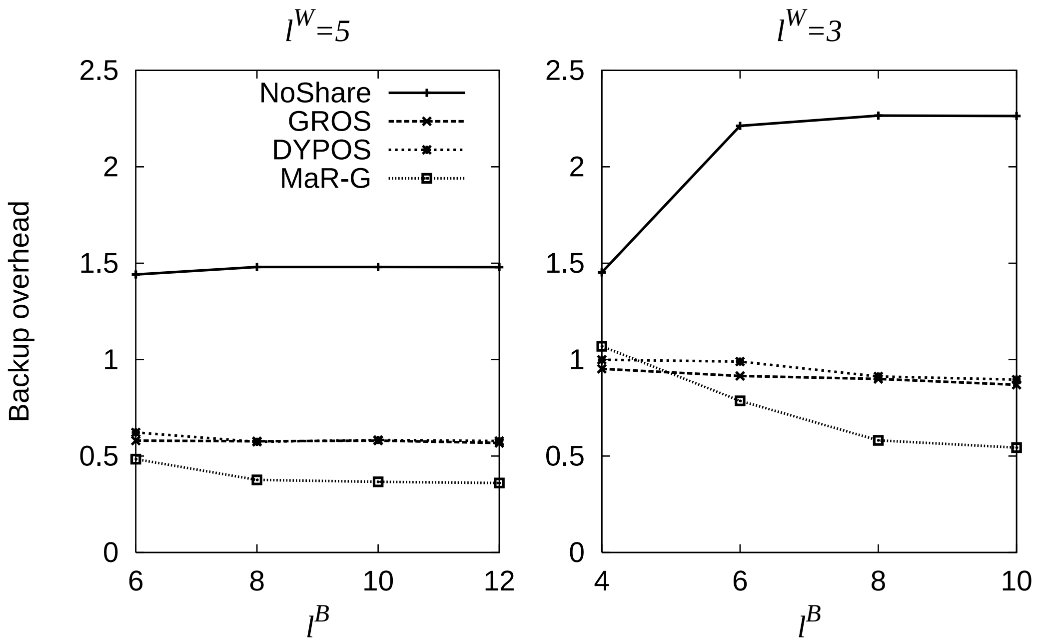

Figure 13: Backup overhead in LARGE-8

6.3.4. Number of sharing cases

Fig. 16 and Fig. 17 show the percentages of requests that benefit from backup bandwidth sharing out of the successfully routed requests in LARGE-5 and LARGE-8. The number of sharing cases are counted under dynamic traffic in order to reflect faithfully bandwidth sharing situation. In all cases, MaR-G encourages more requests to benefit from backup bandwidth sharing than GROS and DYPOS of RaM.

7. Conclusions

In [25] we have proposed the RaM approach for OSSP routing in multi-domain networks. In RaM, approximations are used in cost computing in order to achieve the scalability. In this paper, we have proposed a new approach called MaR where a restricted number of potential intra-paths is selected for carrying traffic across a domain. This restriction allows MaR to benefit from an exact routing which is also highly scalable, although it sacrifices some small possible backup bandwidth sharing and leaves a priori less choices for building working and backup segments. Nevertheless, the

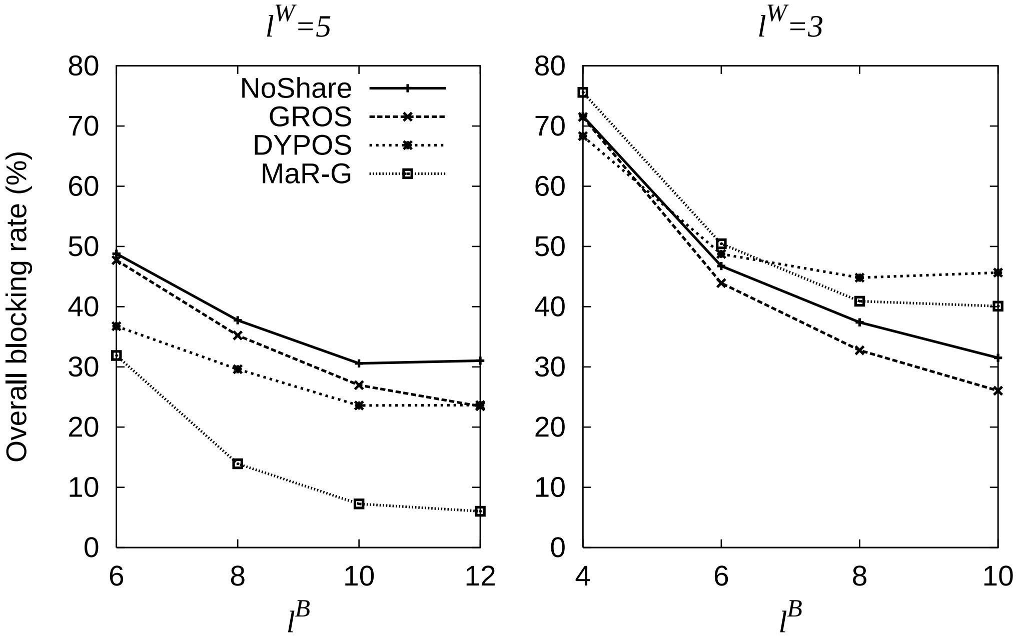

Figure 14: Overall blocking probability in LARGE-5

Mapping with multiple well defined criteria transforms this restriction in a mechanism that orients the Routing to the best intra-paths in terms of cost, disjointness and sharing possibility. In addition, the single routing step of MaR allows improving the quality of bandwidth optimization over the two routing steps of RaM.

The experimental results also confirm that MaR outperforms RaM on bandwidth saving and blocking probability. Furthermore, in bandwidth saving, MaR is close to the optimal single domain solution while the latter is not scalable even for a large single domain network.

MaR can also be applied for WDM multi-domain networks as long as the border node are wavelength conversion capable. Since intra-paths are fixed after the Mapping step, we can allocate statically wavelengths for intra-paths. Wavelengths may need to be changed only at border nodes and each network domain remains all optical without wavelength conversion at internal nodes.

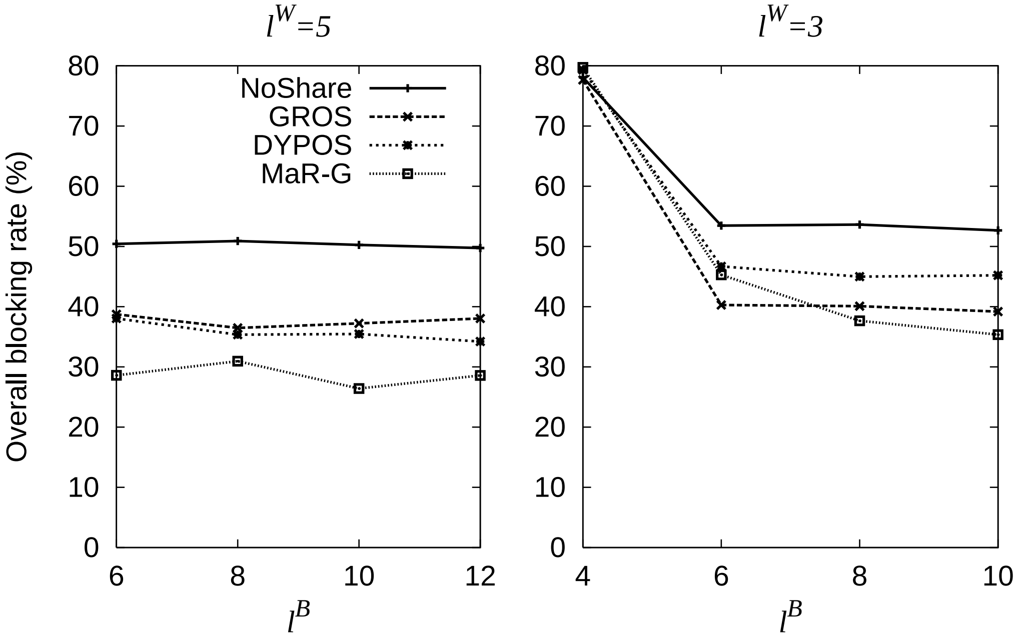

Figure 15: Overall blocking probability in LARGE-8

Acknowledgement

The work of the first author is supported by Vietnams National Foundation for Science and Technology Development (NAFOSTED) under the project number 102.01.13.09.

References

[1] R. Bhandari, Survivable Networks: Algorithms for Diverse Routing, Kluwer Academic, 1999.

[2] P. Datta, A. K. Somani, Graph transformation approaches for diverse routing in shared risk resource group (srrg) failures, Computer networks Journal 52 (12) (2008) 2381–2394.

[3] B. Ramamurthy, S.; Mukherjee, Survivable WDM Mesh Networks, Part I - Protection, in: Proc. IEEE INFOCOM, Vol. 2, New York, NY, 1999, pp. 744–751.

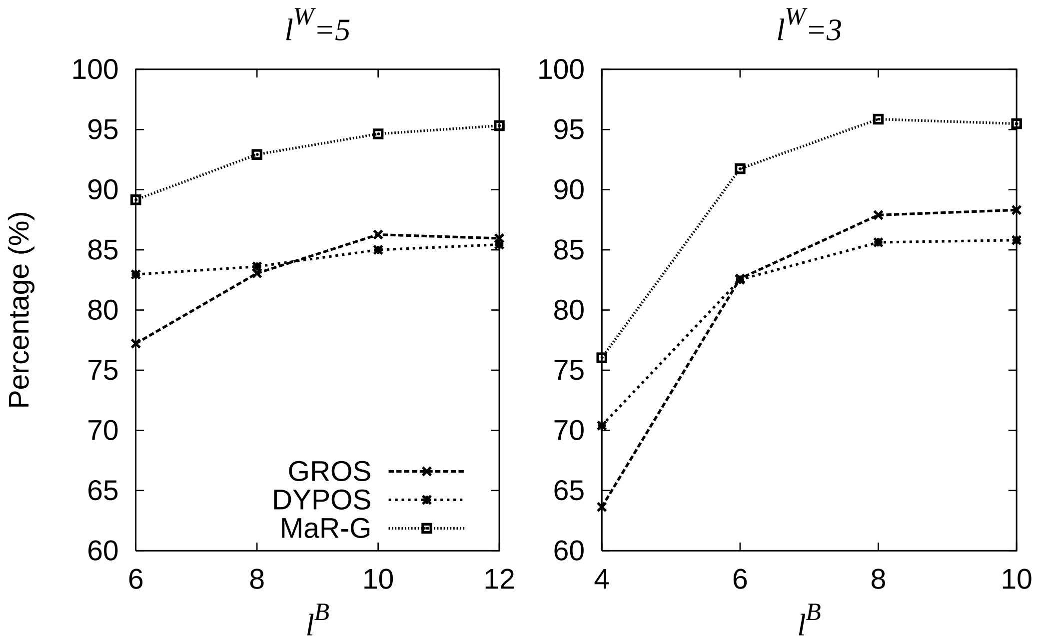

Figure 16: Percentage of the number of bandwidth shared requests over the number of routed requests in LARGE-5

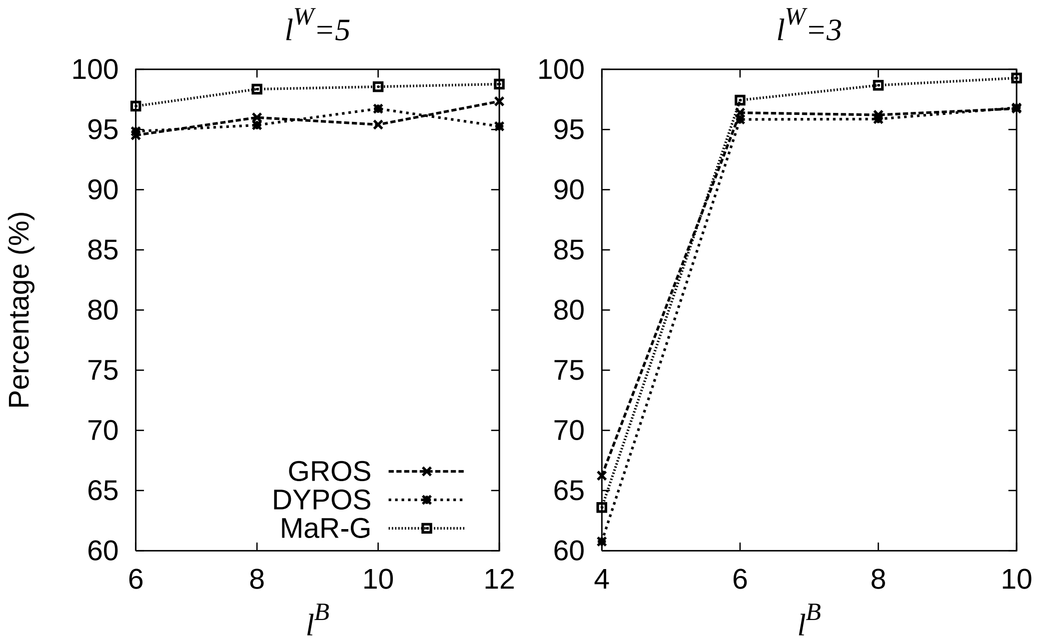

Figure 17: Percentage of the number of bandwidth shared requests over the number of routed requests in LARGE-8

[4] P.-H. Ho, J. Tapolcai, T. Cinkler, Segment shared protection in mesh communications networks with bandwidth guaranteed tunnels, IEEE/ACM Transactions on Networking 12 (6) (2004) 1105–1118. doi:http://dx.doi.org/10.1109/TNET.2004.838592.

[5] W. Grover, D. Stamatelakis, Cycle-Oriented Distributed Preconfiguration: Ring-like Speed with Mesh-like Capacity for Self-planning Network Restoration, in: Proc. IEEE ICC, Atlanta, USA, 1998.

[6] G. Shen, W. D. Grover, Extending the p-Cycle Concept to Path Segment Protection for Span and Node Failure Recovery, IEEE JSAC Optical Communications and Networking 21 (8) (2003) 1306–1319.

[7] G. Shen, W. D. Grover, Segment-based approaches to survivable translucent network design under various ultra-long-haul system reach capabilities, OSA Journal of Optical Networking 3 (1) (2004) 1–24.

[8] M. Kodialam, T. Lakshman, Dynamic Routing of Restorable Bandwidth-Guaranteed Tunnels Using Aggregated Network Resource Usage Information, IEEE/ACM Transactions on Networking 11 (3) (2003) 399–410.

[9] G. Ranjith, G. P. Krishna, C. S. R. Murthy, A distributed primary-segmented backup scheme for dependable real-time communication in multihop networks, in: Proc. International Parallel and Distributed Processing Symposium, 2002, pp. 139–146.

[10] G. Mohan, A. K. Somani, S. Murthi, Efficient algorithms for routing dependable connections in wdm optical networks, IEEE/ACM Transactions On Networking 9 (5) (2001) 553–566.

[11] P.-H. Ho, H. T. Mouftah, Spare capacity allocation for WDM mesh networks with partial wavelength conversion capacity, in: Workshop on High Performance Switching and Routing, 2003, pp. 195–199.

[12] P.-H. Ho, H. T. Mouftah, A framework for service-guaranteed shared protection in WDM mesh networks, IEEE Communications Magazine 40 (2) (2002) 97–103.

[13] D. Xu, Y. Xiong, C. Qiao, Protection with Multi-Segments (PROMISE) in Networks with Shared Risk Link Groups (SRLG), in: Proc. The 40th Annual Allerton Conference on Communication, 2002, pp. 1320–1331.

[14] J. Cao, L. Guo, H. Yu, L. Li, A novel recursive shared segment protection algorithm in survivable WDM networks, Journal of Network and Computer Application 30 (2) (2007) 677–694. doi:http://dx.doi.org/10.1016/j.jnca.2005.12.003.

[15] G. Bernstein, V. Sharma, L. Ong, Interdomain Optical routing, OSA Journal of Optical Networking 1 (2) (2002) 80–92.

[16] J. L. Le Roux, J. P. Vasseur, J. Boyle, Requirements for Inter-area MPLS Traffic Engineering, Tech. rep., IETF Internet-Draft, draft-ietf-tewg-interarea-mpls-te-req-02.txt (Jun. 2004).

[17] D.-L. Truong, B. Jaumard, Multidomain optical networks: issues and challenges - recent progress in dynamic routing for shared protection in multidomain networks, IEEE Communications Magazine 46 (6) (2008) 112–119.

[18] H. Drid, B. Cousin, M. Molnar, S. Lahoud, A survey of survivability in multi-domain optical networks, Comput. Commun. 33 (2010) 1005–1012. doi:http://dx.doi.org/10.1016/j.comcom.2010.02.003.
URL `http://dx.doi.org/10.1016/j.comcom.2010.02.003`

[19] A. Akyamac, S. Sengupta, J.-F. Labourdette, S. Chaudhuri, S. French, Reliability in Single domain vs. Multi domain Optical Mesh Networks, in: Proc. National Fiber Optic Engineers Conference, Dallas, Texas, 2002.

[20] C. Ou, B. Mukherjee, H. Zang, Sub-Path Protection for Scalability and Fast Recovery in WDM Mesh Networks, in: Proc. OSA Optical Fiber Communication Conference (OFC), Vol. 54, Anaheim, California, 2001, p. ThO6.

[21] L. Guo, LSSP: A novel local segment-shared protection for multi-domain optical mesh networks, Computer Communications 30 (8) (2007) 1794–1801.

[22] A. Farkas, J. Szigeti, T. Cinkler, P-cycle Based Protection Schemes for Multi-Domain Networks, in: Proc. International Workshop on Design of Reliable Communication Networks (DRCN), Italy, 2005, pp. 223–230.

[23] H. Drid, S. Lahoud, B. Cousin, M. Molnr, Survivability in multi-domain optical networks using p-cycles, Photonic Network Communications 19 (2010) 81–89, 10.1007/s11107-009-0213-y.
URL http://dx.doi.org/10.1007/s11107-009-0213-y

[24] D. L. Truong, B. Thiongane, Dynamic routing for Shared Path Protection in Multidomain optical mesh networks, OSA Journal of Optical Networking 5 (1) (2006) 58–74.

[25] D. L. Truong, B. Jaumard, Using Topology Aggregation for Efficient Segment Shared Protection Solutions in Multi-domain networks, IEEE Journal of Selected Areas in Communication/ Optical Communications and Networking series 25 (9) (2007) 96–107.

[26] D. Xu, C. Qiao, Y. Xiong, An Ultra-Fast Shared Path Protection Scheme Distributed Partial Information Management, Part II, in: Proc. 10th IEEE International Conference in Network Protocols, France, 2002, pp. 344–353.

[27] B. Rajagopalan, D. Saha, Link bundling in optical networks, Tech. rep., Internet Draft (2001).

[28] CPLEX, http://www.ilog.com/products/cplex/ (2007).

[29] Opnet Modeler, http://www.opnet.com/solutions/network_rd/modeler.html (2007).

[30] M. O'Mahony, D. Simeonidu, A. Yu, J. Zhou, The Design of the European Optical Network, Journal of Ligthwave Technology 13 (5) (1995) 817–828.

[31] REDIrid, http://www.rediris.es/red/index.en.html#red%20troncal (2005).

[32] Consortium GARR, http://www.garr.it (2005).

[33] RENATER-4 network, http://www.renater.fr (2005).

[34] Surfnet, http://www.surfnet.nl (2005).

[35] E. W. Zegura, K. L. Calvert, S. Bhattacharjee, How to Model an Internetwork, in: IEEE Infocom, Vol. 2, San Francisco, CA, 1996, pp. 594–602.

[36] D. Magoni, J. J. Pansiot, Analysis of the autonomous system network topology, SIGCOMM Computer Communication Review 31 (3) (2001) 26–37.